Study of Adherence to Guidelines and Evidence (SAGE):

Theory-based analyses of beliefs, attitudes and prescribing outcomes in British primary care

Arash Rashidian MD

Doctor of Philosophy

University of York

Department of Health Sciences

September 2004

Abstract

The PhD project explored the issues that influenced primary care physician adherence to clinical guideline prescribing recommendations. The project was based on three linked studies that used different methodologies. The first study was a systematic review (overview) of systematic reviews supported by a selective review of theories of behaviour change. It aimed to identify effective methods of improving primary care prescribing. It concluded that multi-faceted interventions were not necessarily more effective than single interventions. The review resulted in taxonomy of interventions to change prescribing behaviour that summarised the effectiveness of different interventions.

The second study was the qualitative study. It was designed and analysed with the help of theoretical insights from the theory of planned behaviour (TPB) and the findings of the reviews. Data were analysed using framework method. The analysis identified seven important themes for implementing clinical guideline prescribing recommendations in primary care: 'credibility of content', 'credibility of source', 'presentation', 'influential people', 'organisational factors', 'disease characteristics' and 'dissemination strategy'. Secondary analysis resulted in a simple model for implementation of guidelines in primary care. The taxonomy of interventions was updated using the findings of the qualitative study.

The qualitative study was exploited for the design of the third study tools (i.e. the surveys' questionnaires). The validity of different methods of sample size calculation for TPB surveys was also assessed. Two stratified random samples of GPs in England were studied. The surveys directly assessed the merits of TPB for understanding GP prescribing. They measured GPs' attitudes and beliefs and their intentions to prescribe according to clinical guidelines for asthma and of statins. Prescribing data were obtained from routine data sources. TPB explained some of the variations in asthma and statins prescribing intentions and behaviours. The surveys demonstrated that GPs views and beliefs contributed to the variations in their prescribing. They also suggested that the effects of GPs' beliefs on their prescribing were not necessarily mediated through their behavioural intentions.

List of Contents

List of Tables

List of Figures

Acknowledgements

This research was completed by the financial support of the Tehran University of Medical Sciences (TUMS) and the Iranian Ministry of Health and Medical Education. I am very grateful for their support.

I am indebted to my supervisor Prof Ian Russell, currently at the University of Wales Bangor, for support, encouragement and advice I received from him throughout this research. He kindly advised me on planning, conducting, analysis, and interpretation of the results and commented on all chapters of the thesis. I am very grateful to my advisor Prof Christine Godfrey, University of York, for her support and advice on the study design, for her comments on different chapters of the thesis. She also supervised final revisions of the thesis, which I found invaluable.

I am very grateful to my advisor Prof Martin Eccles, University of Newcastle Upon Tyne, for his advice on planning of the qualitative study and the surveys, and for his advice on the analysis and interpretation of the results. I am grateful to Dr Jeremy Miles, University of York, for his advice on analytical aspects of Chapters 6 and 7 and to Dr Daphne Russell, University of Wales Bangor, for her advice on Chapter 6.

I sincerely thank all those individuals, who I cannot name to observe their confidentiality, who kindly agreed to spend their invaluable time to take part in the interviews or to respond to the questionnaires. Without their time and effort this thesis would have not been possible to conduct.

I am grateful to Dr Jan van der Meulen, London School of Hygiene and Tropical Medicine for his support for the time I spent on the final parts of this thesis while being employed by the LSHTM. I would like to thank him and Prof Nick Mays, LSHTM, for reading different chapters of the thesis and providing their valuable feedback. The thesis also benefited from the comments raised by the examiners Prof Gene Feder, Queen Mary University of London and Dr Dawn Dowding, University of York and I am grateful to them.

Following individuals helped me with queries and questions at different stages for which I thank them: Dr Hilary Bekker (the interview schedule and the questionnaires); Janette Colclough (search strategy); Michele Cossiy (use of PACT data); Simon

Coulton (random number generation); Prof Nick Freemantle (earlier plans of the study); Prof Bill Gillespie (support of the project); Prof Hugh Gravelle (study proposal); Prof Steve Harrison (providing the opportunity of presenting in the ESRC's clinical governance seminar series); Ken Headly (provision of access to PACT data); Prof Marie Johnston (study proposal and the proposed interventional study in Appendix VIII); Dr Jane Noyes (the qualitative study); Dr Nigel Rice (the surveys design and Appendix VIII); Prof Trevor Sheldon (support of the project, provision of partial funding for PACT data entry and Appendix VIII); Denise Shingler and Diane Stockdale (support as the departmental and graduate school secretaries); Dr Henry Smithson (qualitative study and the questionnaires); Dr Andrew Wagner and Dr Deborah Baker (access to the National PCT Database and GPs' addresses database); Prof Ian Watt and Prof David Torgerson (Appendix VIII) and Dr Pam Levin, Dr Michael Cabana, Dr David Davis, Dr Pierre Durieux, Dr Susan Millstein and Dr Donald Pathman (queries about their published studies).

I would also like to thank members of the TUMS's School of Public Health scholarship interview panel (including Prof Hossein Malek-Afzali, Prof Ali-Reza Mesdaghinia, Prof Nematipour) for choosing me for the PhD scholarship award; Prof Ali-Reza Mesdaghinia and Prof Mohammad-Reza Zafarghandi for their support; Prof Ali-Reza Mesdaghinia, Prof Bagher Larijani, Dr Behrouz Nabaie and Prof Mahmoud Doosti for their supportive reference letters; and Dr Akbari, Dr Naser Shams and Dr Ali Khodaii and their colleagues for general and administrative supports.

Naeimeh Pakatchi, my beloved wife, gave me emotional and intellectual support throughout the PhD period. I am indebted to her and to my delightful and smiling daughter Laleh Rashidian for their remarkable patience and love throughout the time I worked on this thesis and because of evenings and weekends that I spent on it and not with them. My cute little daughter Leila also responded with laughter and curiosity to the endless periods I spent on my laptop. Naeimeh also kindly proofread a few chapters of the thesis. I cannot thank them enough.

Author's Declaration

The author initiated and conducted all the research and material presented in this thesis. The author performed all data handling, transcribing of qualitative data and qualitative and quantitative analyses reported in this thesis. The design of the project and the sub-studies and the interpretation of the findings followed discussion with Prof Ian Russell. However, the author is responsible for the research presented here.

Arash Rashidian

September 2004

Patient No. 1: Lev Nikolayevitch Tolstoy. Sanguine temperament. His illusion is that he can change others' lives by words

Troyat, 1970; quoted from Horder et al, 1986.

Chapter 1. Introduction to the Study of Adherence to Guidelines and Evidence (SAGE)

The subject of this thesis is to explore the issues that influence primary care physicians in following prescribing guidelines. In this chapter the background to the thesis is explained and an outline of the rest of the thesis is presented.

1.1. The dilemma of implementing new (and not so new) innovations

Health care innovations, including prescribing, usually take a long time before being implemented in practice. It took about fifty years (from 1747 to 1795) for the British navy, and a further 70 years until 1865 for merchant fleet, to use lemon juice for the prevention of scurvy (Haines and Jones, 1994). Arguably nowadays many innovations take much shorter to be implemented, but the expectations are now greater. Policy-makers, patients, managers and general public like to see new innovations being implemented in clinical practice soon after their effectiveness being examined and approved; and those innovations for which there are no evidence of support or are questioned by research to be excluded from practice. Health professionals are keen to implement 'new' advances, but this often means using more expensive interventions with limited advantage or questionable effectiveness. Several examples exist. A study in

1966 indicated that tetracycline, a 'wonder drug', took only seventeen months to be prescribed by all physicians in four communities of Illinois (Rogers, 1995b). Furthermore, it was estimated that as much as a quarter of high technology health services, namely cardiac and vascular surgeries, might not have been required (Borowitz and Sheldon, 1993). The dilemma of evidence-based health care starts here: what are the best ways of helping health professionals keeping abreast of new innovations, while doing it in a 'conscientious, explicit, and judicious' manner (Sackett et al, 1996) i.e. avoiding innovations with questionable effectiveness and efficiency?

1.2. Clinical practice guidelines and behaviour change

Governments as well as pharmaceutical industry and charitable organisations spend a lot of resources on medical and clinical research. In comparison, little is spent on the implementation of available evidence (Eddy, 1982).

> *The profession has placed huge value on developing the basic science of medicine; it has not emphasised the process by which the science is translated into practice* (Eddy, 1982; quoted from Lomas and Haynes, 1988, p 78).

Similarly, little work has been done on understanding how to influence GPs' practice (Horder et al, 1986). Most of behaviour change interventions have been based on the assumption that clinicians would change if they are given information (Soumerai et al, 1989). These assumed models of changes are called 'production-dissemination' (Wood et al, 1998) or 'information deficit' models of behaviour change (Schwartz et al, 1989; Marteau et al, 2002). Relying on these models of change, interventionists endeavoured to improve quality of care by providing more information on safety, efficacy and cost-effectiveness of intended behaviours. The inevitable outcome of this approach has featured in abundance of information delivered to medical practitioners (Hibble et al, 1998). In 1991, Tong predicted that decision makers and clinical practitioners would use consensus reports (guidelines) more and more. Her

reasoning was that clinicians were '*bombarded by information*', faced '*conflicting viewpoints*' and were '*uncertain about what to uphold as a value or recognise as a fact*'.

Several studies have shown that in the era of uncertainty health care providers vary substantially in what they provide (Eddy, 1984). Clinical practice guidelines are sought as tools to reduce variation in health care and to reduce cost (Borowitz and Sheldon, 1993; O'Brien et al, 2000) and more importantly to improve quality of patient care (Feder et al, 1999). Increasingly the research findings are summarised in guidelines and a new industry has appeared concerned with guideline development and implementation (Freemantle et al, 1998). Many clinical guidelines have no clear implementation plans and are mainly intended as tools for information transfer. However, research suggests that doctors may not use guidelines as the main source of information (Timpka et al, 1989). Also the majority of clinical guidelines have not been through rigorous production processes, making it more difficult for clinicians to follow their recommendations (Grilli et al, 2000; Graham et al, 2001). It is even claimed that the quality of guidelines is declining (Hasenfeld and Shekelle, 2003). Clinical guidelines should be viewed as health technologies and their effectiveness should be evaluated as any other technology (Grimshaw and Russell, 1993). Evaluation of guideline implementation programmes involves careful planning and requires dedicated resources. Multi-stage studies using qualitative and quantitative methods have been recommended (Campbell et al, 2000a). Still many believe that the experience in (evaluation of) guideline implementation is limited if not scarce (Wensing et al, 1999).

'*The success of clinical practice guidelines depends ... on their widespread application in routine medical practice*' (Mittman et al, 1992, p 413). If not implemented, guidelines are '*words without action*' (Lomas, 1991), even though guidelines can create awareness and work as '*words that indirectly lead to action*' (Rogers, 1995b). Many guidelines fail the implementation phase. Some scholars raise concerns that clinical guideline implementation is in danger of 'falling' after its rapid rise (Smith, 2000). The question is which interventions are more effective for implementation of guidelines? Historically, the evidence on this has been limited. When trying to investigate the available evidence for

the best methods of influencing GP behaviour, Horder et al (1986) were surprised of the scarcity of evidence. Even where there was evidence, it had rarely been acted upon, resulting in waste of scarce resources. A systematic review in 1992 found very limited support for the effectiveness of continuing medical education programmes (Davis et al, 1992). Nonetheless, in 1994 it was estimated that still around 2000 pounds per GP was spent on continuing medical education (Haines and Jones, 1994).

In recent years there have been a lot of attempts to identify effective ways of using clinical guidelines in behaviour change – see Grimshaw and Russell (1993) as a pioneering example – but important questions remain. For some generic questions, there may never be a concrete answer, e.g. what is the best method to implement guidelines? Although it is argued for long that single strategies are less likely to change clinical practice (Stocking, 1992), this is not much of comfort. Multi-facet interventions are costly and more difficult to implement. Why some guidelines were more successfully implemented than the others? Was it because of the differences in quality of guidelines, setting, clinical condition or dissemination strategies? Previous studies suggested that enthusiastic clinicians or 'innovators' achieved more with poor guidelines than what others achieved with better quality guidelines (North of England Study of Standards and Performance in General Practice, 1992a; North of England Study of Standards and Performance in General Practice, 1992b). Why did some dissemination strategies (e.g. educational outreach visits) work well in influencing prescribing in some settings (Avorn and Soumerai, 1983) but not in other settings (Freemantle et al, 2002)?

1.3. Why study prescribing?

Prescribing is one of the most prominent activities of primary care physicians and other office-based doctors. Soumerai et al (1989) estimated that about 75% of visits to office-based doctors end up in prescribing. Several forces influence physician prescribing. Among those are pharmaceutical companies. It has been

estimated that in the early 1980s about $1.5 billion was spent on drug detailing in the USA (Pippalla et al, 1995). Health systems' resources to influence provider behaviour are easily dwarfed by these figures. This is another reason for furthering attempts for better understanding of provider behaviour and methods to influence it.

The WHO refers to the ideal state of prescribing, distribution and use of drugs as 'rational drug use' and provides this definition:

> *The rational use of drugs requires that patients receive medications appropriate to their clinical needs, in doses that meet their own individual requirements for an adequate period of time, and at the lowest cost to them and their community*; WHO conference of experts Nairobi 1985, cited in (Le Grand et al, 1999; Holloway, 2004).

Inappropriate use of drugs can be the result of a variety of situations. Prescribed drug may not be appropriate for patient need, it may be expensive (for patient or system) or it may not be acceptable for personal, cultural or social reasons. Clinicians may prescribe medicines of no value because of perceived patient pressure or placebo effects. They may also prescribe where medication does not provide any benefit over 'wait and see' approach. Irrational prescribing may also be the result of under prescribing of required medicines. This is more often the case for management of chronic diseases or in primary and secondary prevention practices. Chrischilles and Gondek (1997) suggested that appropriateness of prescribing should be assessed at three levels. The first level is whether any medication is at all warranted (alternatives are non-drug treatment or no treatment). The second level is to establish which drug is preferred based on efficacy, effectiveness and safety. The third level focuses on technical issues of prescribing such as dosage, duration, monitoring and drug interactions (Chrischilles and Gondek, 1997). Although this is a useful categorisation, it should be noted that these three levels are correlated and non-exclusive. For example decision to prescribe is usually linked to the availability of appropriate medicines. Similarly a drug may not be suitable for an individual patient because of dosage, monitoring, duration or interactions. It is also not clear why efficiency, or simply cost, is not included among issues that affect the appropriateness of prescribing.

Irrational use of drugs is not always because of physicians. Appropriately prescribed medicines may be used inappropriately. Patients may not use the specified doses of medicines in appropriate intervals or for prescribed durations. Also the role of dispensers should not be overlooked. In many countries pharmacies are allowed to prescribe a range of drugs without physician (or other clinician) prescription. Over-the-counter prescriptions may play a role in irrational use of drugs. Dispensers may also prescribe drugs that require a prescription without prescription. This phenomenon is not uncommon in many countries (Dinarvand and Nikzad, 2000, Hafeez et al, 2004). It is also part of another problem which is self-medication. In countries where the medicines market is not adequately regulated, patients may decide on what they need and obtain it freely from dispensers. The problem of irrational use of drugs may be the result of system failure. Inadequate financial support for patients with chronic or serious infectious diseases and substantial co-payments put disadvantaged groups in unfavourable situation in terms of access to drugs.

Prescribing costs have been growing 6-8% per year in the global context (Le Grand et al, 1999). In Australia in one year prescribing costs rose more than 23% (Beilby and Silagy, 1997), and in the USA the prescribing expenditure increased thirteen-fold in only thirty years from 1960-1990 (Pippalla et al, 1995). Most prescribing costs happen in primary care. It is estimated that only a quarter of drug expenditure happens within hospitals and the rest are due to office-based activities (Chrischilles and Gondek, 1997). In 1995, GP prescribing amounted to 11% of total NHS spending (Majeed et al, 1997). Therefore, focusing on quality and cost of prescribing in primary care is important and vital. In particular if one considers that inappropriate prescribing often results in significant morbidities for patients and avoidable consequent charges for societies and health systems.

1.4. What is primary care?

In the past thirty years, different definitions have been provided for general practitioner (GP) or primary care physician (PCP) (Table 1.1). These definitions have endeavoured to present the ideal content of general practice and the characteristics that separates it from other medically qualified professions (Olesen et al, 2000).

TABLE 1.1. DEFINITIONS OF GENERAL PRACTITIONER (GP) OR PRIMARY CARE PHYSICIAN (PCP)

Definition	Source
The general practitioner is a licensed medical graduate who gives personal, primary and continuing care to individuals, families and a practice population irrespective of age, sex and illness. It is the synthesis of these functions which is unique	Leeuwenhorst, 1974[1]
Primary care physicians are medically qualified physicians who provide primary health care. Primary health care provides integrated, easy to access, health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained and continuous relationship with patients, and practising in the context of family and community	Institute of Medicine; Vanselow et al, 1995
The general practitioner is a specialist trained to work in the front line of a healthcare system and to take the initial steps to provide care for any health problem(s) that patients may have. The general practitioner takes care of individuals in a society, irrespective of the patient's type of disease or other personal and social characteristics, and organises the resources available in the healthcare system to the best advantage of the patients. The general practitioner engages with autonomous individuals across the fields of prevention, diagnosis, cure, care, and palliation, using and integrating the sciences of biomedicine, medical psychology, and medical sociology	Olesen et al, 2000
Medical health care professionals providing first contact and on-going care to patients, regardless of the patient's age, gender or presenting problem	Bower and Sibbald, 1999

[1] Quoted from Olesen et al, 2000

General practice or primary care is different from other health care settings in terms of its organisation, focus, patient case-mix and patient-doctor relationship. Also the composition of health care team is different. The reaction of GPs to a given condition (e.g. a patient with low back pain) may not be similar to that of a consultant physician, a surgeon or a tertiary care specialist. Despite the definitions (Table 1.1) general practice or primary care is not the same thing in different countries. The level of PCP specialisation is variable. In many countries (e.g. the UK) PCPs have to take some training in primary care (as a resident, registrar or trainee) before they qualify for the job. However in some

other countries (e.g. Iran) PCPs are general physicians. Even in countries with specialised training, PCP may mean different things. In the USA some community office-based general paediatricians or internists are rightly considered as PCPs. In the UK there are many GPs with special interest (e.g. in diabetes, asthma, dermatology), who tend to spend part of their office time practicing it and may receive referrals from their peers. In certain countries (e.g. Canada) different parts of the health system have different primary care structures.

There is 'no magic bullet' for provider behaviour change (Oxman et al, 1995). An effective intervention in some circumstances (e.g. in hospital setting) may be less effective or ineffective in other circumstances (e.g. in primary care). In that sense, the issue of health professional behaviour is not different from health care itself. There is no 'wonder drug' to treat all conditions in all 'patients'. Understanding the behaviour of PCPs requires careful studies in primary care. Selective approaches would help health services researchers to identify the questions for which there are answers and the questions that require further investigations.

1.5. What is the thesis about?

The thesis aimed to explain variation in primary care prescribing in accordance to the best available evidence (i.e. clinical guidelines), using theory-based approaches. The thesis presents the findings of a health services research study. The study benefited from the contribution of a few academic disciplines including health psychology, health policy, epidemiology and biostatistics. More importantly the primary findings were the results of different research methodologies: systematic reviewing of evidence, qualitative analysis of interview data and large-scale mailed surveys. These methodologies were used at different stages of the study. Methods used for this thesis can be found in Chapters 2, 4, 6 and 7. The Theory of Planned Behaviour (TPB) was the major theory used in the thesis. The study started with a wide approach (within

primary care prescribing) and as progressed through the stages focused on prescribing for specific clinical conditions.

The study started with an overview of systematic reviews (Chapter 2). The overview was used to collate the available knowledge in terms of the effectiveness of interventions to improve prescribing in primary care. The overview was not limited to any specified clinical conditions. It covered several interventions likely to influence prescribing. It concluded with an updated and improved taxonomy (the SAGE taxonomy) of interventions summarising what was known about the effectiveness of different interventions. The overview also concluded that the widely held view on the effectiveness of multi-faceted interventions could be misleading.

Chapter 3 reports the findings of a selective review of theories of behaviour change. The theories were derived from different academic disciplines. The chapter was intended to contribute to the understanding of variations observed in the effectiveness of the interventions. The review was used for improving the taxonomy developed in Chapter 2. In this chapter the potential merits of TPB for explaining variation in primary care prescribing were also presented.

Chapter 4 reports the findings of a qualitative study of GPs and academic of primary care in Britain. The qualitative study focused on views and concerns of GPs about factors that contributed to the success (or failure) of clinical guidelines in promoting effective prescribing in primary care. Five clinical conditions were specifically considered in the interviews. These were depression, menorrhagia, statins for coronary heart disease (CHD) prevention, asthma and epilepsy. The TPB was used in the design of the interview plan and in devising the framework for data analysis. Based on the qualitative study two clinical conditions (asthma and statins for CHD prevention) were identified to be used in the surveys. The qualitative analysis resulted in identifying seven main themes (and 30 sub-themes) for implementation of prescribing recommendations. Not all the themes were relevant to interventions, but using those which were, the taxonomy of intervention was further developed. Based on secondary analysis of the qualitative data and the thematic framework, a

simple model for the implementation of clinical guidelines for prescribing in British primary care was developed (Chapter 5).

Before conducting the surveys, it was noted that previous TPB studies had mostly shunned the issue of sample size estimation. Methodological analyses were conducted to establish the optimal methods of sample size calculation for a TPB study (Chapter 6). The findings could be useful for sample size calculation of any study using linear regression analysis. The results of two national surveys of GPs intentions to prescribe in accordance with clinical guidelines for asthma and prescribe statins for the prevention of CHD are reported in Chapter 7. The surveys were based on TPB. The questionnaires were developed based on the findings of the qualitative study (Chapter 4). The surveys examined the theory as if it provided better understanding of GPs' variation in attitudes towards, and intentions to use clinical guidelines. The surveys also assessed TPB's ability in explaining prescribing variation, using routinely collected dispensing data.

Chapter 8 is the final chapter of the thesis. In this chapter a summary of the findings and important discussions and implications as well as main limitations were provided. It concluded that the TPB had important potentials for guideline implementation and GP prescribing. It also summarised the main policy and research implications of the thesis.

Belief in education as a method of influencing general practitioners is confirmed. But gains in knowledge, skills and changes in behaviour seem harder to achieve.

Horder, Bosanquet and Stocking, 1986

Chapter 2: Overview of systematic reviews. Effective interventions for improving physician prescribing in primary care

2.1. Background

This chapter presents the results of an overview (i.e. systematic review) of systematic reviews of health professional behaviour change interventions with focus on physician prescribing in primary care (i.e. PCP prescribing). The results of the review in terms of the effectiveness of different interventions for improving PCP prescribing are presented. The chapter ends with a table (Table 2.1) summarising and categorising the findings in the taxonomy of interventions.

Implementation of clinical guidelines and changing physician behaviour has proved difficult. Within the last two decade several systematic reviews of provider behaviour change have been published. These reviews attempted to summarise the state of the art and knowledge by identifying the effectiveness (and sometimes efficiency) of specific interventions to improve provider behaviour. In turn, the results of the systematic reviews have been summarised in few overviews of systematic reviews (Conroy and Shannon, 1995; Bero et al, 1998; NHS Centre for Reviews and Dissemination, 1999; Cantillon and Jones, 1999; Durieux et al, 2000; Smith, 2000; Grimshaw et al, 2001).

One earlier and frequently cited overview focused on the effectiveness of different interventions to improve the implementation of research evidence (Bero et al, 1998). Eighteen eligible systematic reviews were included in that study. Bero et al found that many primary studies included in systematic reviews were methodologically flawed. Given the observed effect sizes were small, this might have resulted in erroneous conclusions. They also observed few studies performing economic evaluation. Another shortcoming identified by Bero et al was lack of generalisability. Many primary studies had originated from North America. Even in their North American context most studies had been conducted by small number of researchers in limited range of settings. Nonetheless the overview had important conclusions; mainly it discouraged use of passive dissemination of educational materials. This conclusion has been repeated in several studies published since Bero et al (1998).

Another widely disseminated overview of systematic reviews was published as an Effective Health Care bulletin (NHS Centre for Reviews and Dissemination, 1999). It updated Bero et al (1998) study. Forty-four studies were included in this update. NHS CRD (1999) was effectively more than an overview of systematic reviews. It included a review of a number of theories of behaviour change, and addressed different practical and organisational issues that might inhibit evidence implementation. These latter sections of the overview did not follow a systematic approach. Hence NHS CRD (1999) was an authoritative publication for enhancement of evidence implementation and provider behaviour change. This was likely to have improved the relevance and usefulness of the report, but might have negatively affected its reliability. The report highlighted six recommendations on the front page. The level of evidence behind these recommendations varied, but this was not acknowledged. The longer version of NHS CRD (1999) appeared two years later as Medical Care supplement (Grimshaw et al, 2001). Forty-one systematic reviews were included in this later publication. Like its predecessor, it covered the period until mid 1998. The reviews were from a dispersed range of 27 medical journals. The authors found high variability in quality of the systematic reviews. They also found the meta-analyses performed in nine reviews were inappropriate because of high levels of heterogeneity and unit of error analyses in original papers (Grimshaw et al, 2001).They discussed the findings under three categories of

'systematic reviews of broad strategies', 'systematic reviews of specific interventions' and 'systematic reviews of interventions for specific behaviours'. The study had some conclusions that were relevant to prescribing. It concluded that educational outreach visits were effective in changing prescribing. Grimshaw et al (2001) also concluded that audit had variable effectiveness. They also concluded that multi-faceted interventions were more likely to be effective and invited researchers to try to disentangle these in order to identify the effective elements within multi-faceted approaches. They also sought further studies with formal economic evaluation. The main conclusions of this overview were in essence similar to the earlier Bero et al (1998) study.

One overview was published in French (Durieux et al, 2000). The overview focused on using clinical guidelines for behaviour change. It did not discuss prescribing behaviour as a separate outcome (personal correspondence with Pierre Durieux). Other overviews of systematic reviews employed less systematic approaches. One earlier study focused on implementation of clinical guidelines in primary care, but did not pursue a systematic approach and there was little focus on prescribing (Conroy and Shannon, 1995). A later study focused on the effectiveness of continuing medical education in general practice (Cantillon and Jones, 1999). Primary studies as well as systematic reviews were included in this overview. The reporting of the findings did not follow a systematic approach and studies which were unlikely to be systematic reviews or high quality primary research were reported in the result section of the paper. The study had limited findings from interventional studies that were relevant to prescribing. One important finding was that the effects of educational interventions might be short-lived (Cantillon and Jones, 1999).

Smith (2000) reviewed meta-analyses and 'structured' reviews of interventions to change physician behaviour. He also reviewed a number of theories of behaviour change and included primary studies as well as systematic reviews. However, it was unclear why for some interventions primary studies were considered and for other interventions they were not. Perhaps the author intended to fill in the evidence gap where there were no systematic reviews, but the methods of identifying the gaps were not explained. The two

major findings of the overview were similar to those reported by others: no intervention was effective in all circumstances, and multi-faceted approaches were more likely to succeed in behaviour change (Smith, 2000). The overview had a reasonable focus on prescribing but did not distinguish between prescribing in primary care and other settings. It reported that educational outreach, audit and feedback and clinical guidelines were effective in changing prescribing in selected situations. Smith concluded '*in the language of clinical medicine, we must diagnose the lesion (why change is not adopted) before prescribing therapy (a change strategy)*' (p 16S). Lack of an explicit systematic approach in the latter three overviews made their findings more prone to biases because of selective reporting (Conroy and Shannon, 1995; Cantillon and Jones, 1999; Smith, 2000).

None of the aforementioned overviews of systematic reviews focused explicitly on prescribing. The need for more focused approaches to provider behaviour change is well established (Grimshaw and Russell, 1993; Oxman et al, 1995). In line with the objectives of the thesis, this chapter presents the findings of the overview of systematic reviews of provider behaviour change for PCP prescribing (Mulrow, 1994). The overview covered all interventions to improve prescribing regardless of whether they used clinical guidelines or not to ensure all relevant studies were included. The results were also used to update and improve taxonomies of interventions for prescribing behaviour change. Therefore, the objectives of the overview were summarised as follows:

2.1.1. Objectives

The overview had two main objectives:

- To evaluate the effectiveness of interventions to improve primary care prescribing.
- To provide updated taxonomy of interventions for improving primary care prescribing.

2.2. Methods

2.2.1. Inclusion and exclusion criteria

Time frame

- Systematic reviews published from 1980 to 2001 were considered eligible for inclusion.
- The existence of any update for the included Cochrane reviews was checked in 2004 (Cochrane Database, Issue 2, Chichester, UK: John Wiley & Sons).

Study designs

- Systematic reviews of interventional studies were included.
- Systematic reviews that included (cluster) randomised controlled trials (incl. balanced incomplete block, randomised crossover, and simple randomised trials), interrupted time series (ITS) and controlled before-and-after studies (controlled by other doctors or untargeted behaviour, switchback designs) were deemed eligible (Grimshaw and Russell, 1993).
- Included systematic reviews could have considered other types of studies, but only evidence from aforementioned study designs was eligible for inclusion in the overview.
- Reviews of the methodological quality of published studies, studies without explicit research methodology and those of bibliographic nature for published or ongoing research were not eligible (Bero et al, 1998).
- Earlier versions of updated systematic reviews were excluded unless the updates had a shift in focus or incomplete reporting.
- Other study designs were discussed but not 'included' in the overview. These studies were valuable in discussing the effectiveness of interventions for which there was limited evidence in systematic reviews.

Participants

- Systematic reviews that did not exclude primary studies of PCP prescribing were eligible.
- PCPs were defined as all physicians working in primary care that fulfilled the aforementioned definitions (Table 1.1) and were inclusive of GPs, family doctors, family physicians, family practitioners and family medicine specialists.
- In this overview, the definition excluded physicians in training (i.e. students, residents, trainees).
- Physician activities in nursing homes or community hospitals only were also excluded.
- Included systematic reviews could have considered health care providers other than PCPs, but only evidence from studies of PCPs was eligible for inclusion.

Interventions

- Any strategy or combination of strategies to promote effective or efficient prescribing in primary care was considered.
- The strategies included, but were not limited to, continuous medical education (CME), audit and feedback, peer review, reminders, quality improvement cycles (continuous quality improvement, total quality management), clinical guideline implementation programmes, computerised systems, patient education, educational outreach visits, local opinion leaders, financial incentives, remuneration schemes, non-monetary rewards or penalties, interventions involving change in the formulary, mailouts and mass media campaigns and inter-professional strategies. The definitions used for different interventions are provided in the Glossary.

Outcome measures

- It was expected that included systematic reviews would report a wide range of outcomes. The focus of the overview was on prescribing. All outcomes relevant to primary care prescribing were considered, including data recording, prescribing, cost and patient outcomes.

- Data from test environments (role playing, paper patients, vignettes, .simulation cases etc.) were excluded.
- Papers were excluded if the only relevant outcomes were prescribing knowledge or attitude.
- Vaccination, exercise prescribing and investigatory prescribing (i.e. laboratory tests, radiology etc) were excluded.

As a working definition, systematic reviews of the effectiveness of provider behaviour change interventions which did not exclude primary care physician prescribing were eligible for inclusion in the overview of systematic reviews. The included studies were required to meet the previously stated inclusion and exclusion criteria.

2.2.2. Search strategy

Given the broad perspective of the potentially relevant systematic reviews, different approaches were followed to ensure eligible systematic reviews were identified. Several alternative search strategies were devised and tested. In the end it was decided to use a search strategy previously tested for a broad perspective overview of systematic reviews (Grimshaw et al, 2001). The search combined a validated systematic review search with a search for provider behaviour change interventions (Appendix II-1). The search was not limited to any setting (e.g. primary care), provider (e.g. PCPs), behaviour (e.g. prescribing) or language; therefore, it was sensitive enough to ensure potentially eligible studies were not missed. It was used to search Medline and Embase (1980 – 2001) with slight adaptation. The search was complemented using following approaches. Forward citation searches for three broadly cited systematic reviews (Grimshaw and Russell, 1993; Oxman et al, 1995; Davis et al, 1995) were conducted in Science Citation Index and Social Science Citation Index databases (provided by Web of Knowledge – then Web of Science, MIMAS, University of Manchester). The Cochrane Library (Chichester: John Wiley & Sons) – inclusive of DARE (Database of Abstracts of Reviews of Effectiveness, CRD, University of York) and HTA (Health Technology

Assessment) – were also searched. The Cochrane database search was updated in 2004 to consider any potential updates of already included Cochrane reviews. References already included in the author's personal database and those known by his advisers were also considered. The bibliographic references of the included studies, previous overviews of systematic reviews and relevant editorials were also searched. Bibliographic search was iterative and was repeated for every identified literature. In total about twelve thousand titles were considered. All studies that appeared to be reviews of provider behaviour change interventions were flagged for more detailed assessment. After careful consideration against the inclusion criteria, the full-texts of 109 papers that appeared to be reviews of behaviour change interventions were ordered. Out of those, seventy-two reviews were excluded from the overview. List of the excluded studies and the reasons for exclusion are provided in Appendix II-3.

2.2.3. Methods of the review

Definitions, criteria for inclusion and data extraction strategy were agreed upon. The titles and abstracts of identified studies were scanned before ordering the papers. All papers likely to be eligible were ordered. The papers were excluded according to the criteria. After the first round of data extraction, the excluded and included papers were checked again and the extracted data was re-examined. Any doubt or discrepancy was solved by reference to the original papers. When necessary the abstracts or full-texts of primary studies included in systematic reviews were considered. Data was extracted from the systematic reviews; and exceptionally from primary studies in case there was any uncertainty. Primary study studies were also checked to ensure the eligibility for inclusion (e.g. meeting the overview's inclusion criteria in terms of setting, provider, outcome behaviour and study design) and the accuracy or completeness of data reported in the systematic reviews. Data were analysed using narrative synthesis. The effectiveness of single strategies and different combinations of strategies were noted in the analysis. The generalisability of the systematic reviews' conclusions to PCP prescribing was noted. It was decided a priori to compare discordant systematic reviews following a published algorithm

(Jadad et al, 1997). The algorithm was meant for comparing reviews with discordant results, but not reviews with discordant interpretation of results. Jadad et al (1997) suggested that if the results of two or more systematic reviews were discordant, certain steps should be followed. The steps included comparing review questions, included studies, inclusion criteria, review quality and data extraction and analytical strategies.

Quality of systematic reviews was assessed by taking notes of different methodological aspects of the reviews, including search strategies, inclusion criteria, data synthesis strategies, data extraction and whether the reviews' conclusions were supported by the data (Oxman, 1994). No attempts were made to 'formally' assign quality scores to the systematic reviews. Apparent shortcoming in quality and methods of systematic reviews were stated in the evidence table (Appendix II-2). The shortcomings were identified by checking the primary studies included in the reviews (whenever discrepancies were observed) as well as the methods and reporting of the systematic reviews.

2.3. Findings

2.3.1. Description of systematic reviews

Thirty-three systematic reviews (39 papers) that met the inclusion criteria were located (Lomas and Haynes, 1988; Soumerai et al, 1989; Mugford et al, 1991; NHS Centre for Reviews and Dissemination, 1994; Wensing and Grol, 1994; Grimshaw et al, 1995; Pippalla et al, 1995; Yano et al, 1995; Anderson and Lexchin, 1996; Balas et al, 1996; Beilby and Silagy, 1997; Chrischilles and Gondek, 1997; Freemantle et al, 1997; Thomson O'Brien et al, 1997; Hunt et al, 1998; Shortell et al, 1998; Wensing et al, 1998; Bower and Sibbald, 1999; Gill et al, 1999; Giuffrida et al, 1999; Le Grand et al, 1999; Thomson O'Brien et al, 1999; Walton et al, 1999; Zwarenstein et al, 1999; Bower and Sibbald, 2000; Chaix-Couturier et al, 2000; Gosden et al, 2000; Zwarenstein and Bryant, 2000; Zwarenstein et al, 2000; Armour et al, 2001; Gosden et al, 2001; Gross and Pujat, 2001; Lewin et al, 2001; Mitchell and Sullivan, 2001; Ratanawijitrasin et al, 2001; Thomson O'Brien et al, 2001; Walton et al, 2001; Grilli et al, 2002; Jamtvedt et al, 2003).

The included reviews updated eleven previous systematic reviews (Haynes et al, 1984; Haynes and Walker, 1987; Raisch, 1990a; Raisch, 1990b; Davis et al, 1992; Grimshaw and Russell, 1993; Johnston et al, 1994; Oxman et al, 1995; Sullivan and Mitchell, 1995; Davis et al, 1999; Gosden et al, 1999). The majority of the included studies had been published as Cochrane reviews (13 papers). One Cochrane review had been withdrawn, but it was included in the overview since there were no update or substituting review (Freemantle et al, 1997). One review was published in a bulletin and the rest in 18 different journals. Journals were in different categories of general and internal medicine (9 papers), health services and quality of care (6 papers), pharmacy (3 papers), preventive care and social medicine (2 papers) and general practice (1 paper). The included reviews' first authors were based in ten different countries: the USA and the UK (9 reviews each), Canada (8 reviews), Australia, France, Italy,

Kenya, Norway, Thailand and the Netherlands (one review each). The country of location was assessed based on the first authors' affiliations as provided on the papers.

The studies had different points of focus. In terms of the target behaviour, nine reviews focused on prescribing (Soumerai et al, 1989; Pippalla et al, 1995; Anderson and Lexchin, 1996; Chrischilles and Gondek, 1997; Gill et al, 1999; Le Grand et al, 1999; Walton et al, 1999; Ratanawijitrasin et al, 2001; Gross and Pujat, 2001; Walton et al, 2001), one on provision of mental health care (Bower and Sibbald, 2000; Bower and Sibbald, 1999) and one on patient-centred care (Lewin et al, 2001). In the latter review the impact of interventions to improve patient-centred care on provider behaviour (including prescribing) was assessed. Other reviews were not limited to any specific behaviour. Many reviews focused on the effectiveness of specific interventions to improve provider behaviour. Of those, three reviews focused on different types of audit and feedback (Mugford et al, 1991; Beilby and Silagy, 1997; Jamtvedt et al, 2003), four reviews considered financial incentives (Chaix-Couturier et al, 2000; Gosden et al, 2001; Armour et al, 2001; Gosden et al, 2000; Giuffrida et al, 1999) and three focused on computerised systems (Hunt et al, 1998; Walton et al, 1999; Mitchell and Sullivan, 2001; Walton et al, 2001). For each of the following interventions one systematic review was included: mass media interventions (Grilli et al, 2002), doctor-nurse relationship (Zwarenstein and Bryant, 2000), inter-professional education (Zwarenstein et al, 1999; Zwarenstein et al, 2000), substitution or consultation-liaison models of mental health care (Bower and Sibbald, 2000; Bower and Sibbald, 1999), local opinion leaders (Thomson O'Brien et al, 1999), educational outreach (Thomson O'Brien et al, 1997), printed educational material (Freemantle et al, 1997), CQI (Shortell et al, 1998), CME (Thomson O'Brien et al, 2001), physician profiling and peer review (Balas et al, 1996) and national drug policies in less developed countries (Ratanawijitrasin et al, 2001). Five systematic reviews had wider appeals and covered different interventions for different sets of behaviours (Lomas and Haynes, 1988; NHS Centre for Reviews and Dissemination, 1994; Wensing and Grol, 1994; Grimshaw et al, 1995; Yano et al, 1995; Wensing et al, 1998; Gross and Pujat, 2001), of which three focused on implementation of clinical guidelines (NHS Centre for Reviews and Dissemination, 1994; Wensing and

Grol, 1994; Grimshaw et al, 1995; Wensing et al, 1998; Gross and Pujat, 2001). All but nine reviews were not limited to primary care and considered data from different health care settings (Soumerai et al, 1989; Wensing and Grol, 1994; Yano et al, 1995; Anderson and Lexchin, 1996; Beilby and Silagy, 1997; Wensing et al, 1998; Bower and Sibbald, 2000; Mitchell and Sullivan, 2001; Gosden et al, 2001; Gosden et al, 2000; Giuffrida et al, 1999; Bower and Sibbald, 1999). All but two reviews were not limited to any set of countries or regions of the world. Both studies considered evidence from less developed countries only (Le Grand et al, 1999; Ratanawijitrasin et al, 2001). Some reviews although were not confined to a given country, but were more tuned to the needs of the health system of that country, e.g. Canada (Anderson and Lexchin, 1996) and the USA (Shortell et al, 1998).

Many primary studies were included in more than one systematic review. The systematic reviews included 1173 counts of primary studies between themselves, of which 165 counts of papers (representing about 80 unique studies) met the overview's inclusion criteria (i.e. appropriately designed interventional studies of improving PCP prescribing). About a quarter of those 80 studies were included in more than one review, with some being included in more than six systematic reviews (e.g. Avorn and Soumerai, 1983; Schaffner et al, 1983), hence the total of 165. Details of all included systematic reviews and their corresponding primary studies are presented in the evidence table (Appendix II-2). Several Cochrane reviews had not been updated for periods of up to four years (in 2001), despite the Cochrane Collaboration's objective of updating reviews every 1-2 years. Many were not still updated in summer 2004 increasing the delay to up to seven years.

2.3.2. Quality of systematic reviews

As noted in previous overviews of systematic reviews (Bero et al, 1998; Grimshaw et al, 2001) the quality of systematic reviews varied greatly. Attention to the included primary studies was informing since other threats that could invalidate the results of systematic reviews were identified. Some examples of

important quality issues are presented in here and further details are reported in Appendix II-2.

The majority of the reviews synthesised the data using narrative synthesis techniques, although the methods were often left unexplained. Underlying variations in the primary studies is the main barrier to quantitative synthesis of data in systematic reviews of provider behaviour. Biases due to low transparency of the analytical method could not be excluded. The methodology of narrative synthesis is still developing. The UK ESRC Methods Programme is providing guidance for undertaking 'narrative synthesis' which may help standardisation of the technique (Economic & Social Research Council, 2003). One review conducted random effects model meta-analysis based on the standardised mean differences of the outcomes (Walton et al, 1999; Walton et al, 2001). Two systematic reviews used 'vote counting' and 'z transformation' methods and performed quantitative meta-analyses (Balas et al, 1996; Gill et al, 1999). Both reviews acknowledged that there was a great degree of heterogeneity in the data, which undermined the validity of the quantitative approaches. A further review used meta-analysis while the validity of using this approach was questionable (Pippalla et al, 1995).

Three reviews went beyond their stated objectives. Two separate systematic reviews of prescribing included non-prescribing primary studies. One study included primary studies of laboratory test ordering, while it was aimed at assessing the interventions to influence prescribing (Pippalla et al, 1995). The same review also included a totally irrelevant non-interventional study (see Appendix II-2 for details). Another review included studies of preventive behaviour, while its objective was to assess the impacts of clinical guidelines in drug utilisation reviews (Chrischilles and Gondek, 1997). A further review included a study with strong educational outreach component as a study of provision of costing information which might be misleading (Beilby and Silagy, 1997).

The search strategies of a few reviews, including two Cochrane reviews, seemed to have important shortcomings as they missed relevant primary studies included in other reviews (Wensing and Grol, 1994; Pippalla et al, 1995;

Beilby and Silagy, 1997; Chrischilles and Gondek, 1997; Shortell et al, 1998; Freemantle et al, 1997; Gross and Pujat, 2001; Jamtvedt et al, 2003). There were also occasional misses of duplicate publications of primary studies in the systematic reviews (Appendix II-2).

There were errors in classifications of primary studies in terms of their designs or approaches, which might have hampered the validity of the systematic reviews' conclusions. An interesting example was obtained by comparing two reviews of studies from developing countries (Le Grand et al, 1999; Ratanawijitrasin et al, 2001). Both reviews included an earlier study of the effectiveness of essential drugs programmes in Yemen (Hogerzeil et al, 1989; Walker et al, 1990). Le Grand et al (1999) referred to the study as a randomised trial, while Ratanawijitrasin et al (2001) identified the same research as a post-only controlled trial. Checking the original papers of the primary study revealed that it was a post-only controlled trial. In another example a primary study (Lobach and Hammond, 1994) was considered as prescribing study in Chrischilles and Gondek (1997) and not as such in Mitchell and Sullivan (2001). Chrischilles and Gondek included evidence from primary studies without reported prescribing outcomes and used evidence from those studies to draw conclusions for prescribing. One of the systematic reviews wrongly classified two RCTs as non-randomised and one CBA as a randomised study (Le Grand et al, 1999). Similarly, another review did not notify an RCT as such (Chrischilles and Gondek, 1997). A review with quantitative analysis included three separately published analyses of one intervention in the quantitative synthesis as separate studies (Pippalla et al, 1995). Similarly a CBA study was considered as RCT in Balas et al (1996) review.

The reviews varied in the definitions they used for the interventions. This was sometimes because of using non-standard definitions (Lomas and Haynes, 1988) or because of joining some interventions together in one category in order to simplify the analysis processes (Gill et al, 1999).

In total twenty-three interventions were identified in the literature. The following sub-sections of the chapter report the identified evidence for the effectiveness of different interventions. Competence oriented interventions are

introduced first. Then performance oriented and social influence interventions are discussed. Next the interventions involving physical support, financial incentives and non-voluntary strategies are presented. Finally a distinct subsection discusses the evidence for the effectiveness of multi-faceted interventions. Further information about all included systematic reviews and important discussions about the implications or methods of the reviews are included in the evidence table (Appendix II-2; as the table is too long to be included in the text).

2.3.3. Continuing medical education (CME) and inter-professional education

Educational interventions are among the most common interventions to improve provider behaviour. The effectiveness of didactic CME, interactive CME and inter-professional education is discussed in here.

A Cochrane review assessed the effectiveness of CME meetings and workshop in improving provider performance and patient outcomes (Thomson O'Brien et al, 2001). The review was an update of a series of previous reviews (Davis et al, 1992; Davis et al, 1995; Davis et al, 1999). Older studies had wider inclusion criteria and included interventions other than CME (Davis et al, 1992; Davis et al, 1995). The Cochrane review in total included 32 studies of which 30 were RCTs. The review concluded that interactive workshops could result in moderately large changes in provider performance, while didactic sessions were generally ineffective. Four included RCTs focused on prescribing in primary care. An RCT of didactic approach in Sri Lanka suggested no additional effects from educational seminar over and above the effects resulted from an educational newsletter in reducing antibiotic injection rates (Angunawela et al, 1991). Workshops of interactional group discussions were evaluated against no intervention in another RCT conducted in Indonesia. The intervention had moderate to large effects in reducing intramuscular injection rate (Hadiyono et al, 1996). Mixed methods including three educational seminars were used in an RCT in Zambia to improve prescribing, with small but statistically significant

effects on average number of drugs per prescription (Bexell et al, 1996). An RCT of theory based educational seminars for general paediatricians in community practices improved asthmatic patients' outcomes and practitioners' prescribing (Clark et al, 1998). The outcomes of the primary care prescribing studies supported the general conclusion of the review.

A further study was included in a systematic review of clinical guideline implementation in primary care (Wensing and Grol, 1994; Wensing et al, 1998). This was a CBA study of tutorial sessions reporting improvements in physician antibiotic prescribing in the intervention group (Klein et al, 1981).

There was no evidence of effect for inter-professional education. A Cochrane review (Zwarenstein et al, 1999; Zwarenstein et al, 2000) failed to include even a single study on the effects of inter-professional education on process or patient outcomes. The review was not limited to any specific outcome (prescribing or other outcomes). Lack of evidence may be due to the complexity of devising the intervention and evaluating it, although it is possible to plan such intervention. It may also be the result of professionals and others perceiving no equipoise in here. Studies are required to assess the effects of these interventions.

2.3.4. Mailed printed educational material and mailed national warning campaigns

Several systematic reviews discussed the effectiveness of mailed printed educational materials. A Cochrane review assessed the effectiveness of printed educational materials in changing provider performance and patient outcomes (Freemantle et al, 1997). They considered primary studies regardless of the method of the delivery of printed material including mass mailing, delivery by hand or personal mailing. The review included eleven studies of which two targeted prescribing in primary care. In both studies mailed educational materials resulted in small non-significant reductions in inappropriate prescribing. In a Dutch RCT the intervention's effect on undesirable

antispasmodics was very modest (Denig et al, 1990). In the second RCT inappropriate prescribing reduced by 4% (reported 3% in Freemantle et al, 1997) in the group receiving mailed educational material (Avorn and Soumerai, 1983). Freemantle and colleagues (1997) pointed out that the small changes observed from the intervention might have been cost-effective. The reviewers also questioned the evidence base of educational material provided in these two studies, suggesting doubtful evidence base might have contributed to the small effects. In the end the authors concluded that the value of printed material in comparison to no intervention was uncertain.

Another systematic review of prescribing in primary care, Soumerai et al (1989), included two studies from the USA assessing the effects of mailed printed educational material. One RCT was mentioned earlier in this section (Avorn and Soumerai, 1983). The investigators performed an economic evaluation on the results of the trial and concluded that mailed educational material intervention was cost-effective (Soumerai and Avorn, 1986). The review also reported the results of a CBA study in which attractively designed mail brochures did not reduce inappropriate antibiotic prescribing (Schaffner et al, 1983). Soumerai et al (1989) concluded that the evidence did support the effectiveness of the intervention, but they pointed out that it might be cost-effective. A further systematic review (Lomas and Haynes, 1988) included Avorn and Soumerai (1983) as well as another RCT that reported no significant effects from weekly educational materials on the management of hypertension (Evans et al, 1986). The review then concluded '*the resources currently being spent on the production and distribution of such material* [printed educational material] *can be diverted into more effective approaches*' (Lomas and Haynes, 1988, p 87). A more recent systematic review also considered both studies included in Soumerai et al (1989) and concluded that educational material mailouts seemed to be ineffective (Gross and Pujat, 2001).

Similarly Gill et al (1999) concluded that printed educational materials were less effective than other interventions in improving prescribing. This conclusion was based on evidence from a mixture of different health care settings. However, for this overview attention to the relevant subgroup of included studies (i.e. from primary care) revealed that in 2 out of 4 included

studies printed educational material significantly improved antibiotic prescribing. Le Grand et al (1999) review included a three-arm clustered RCT from Sri Lanka. The 'newsletter only' campaign resulted in a reduction of 7.4% in antibiotic prescribing (Angunawela et al, 1991). This reduction was similar to the reduction observed in the 'newsletter plus seminar' group and was further than the change in the control group (0.4%). However, the differences were not statistically significant probably owing to the lack of power (15 health centres, 45 prescribers).

A systematic review of prescribing in primary care identified five ITS studies assessing the effects of mailed national warning campaigns on prescribing (Soumerai et al, 1989). Three European studies (two from Britain and one from Sweden) demonstrated important and statistically significant reductions in the use of chloramphenicol (Wade and Hooh, 1972), pressurised aerosols for asthmatic patients (Inman and Adelstein, 1969) and dipyrone (Bottiger and Westerholm, 1973). Studies also reported reductions in the rate of associated serious adverse events. In all studies mailed national warnings to doctors were accompanied by publication of letters and articles in professional journals. Fourth study had been conducted in the USA and failed to demonstrate any reduction in the use of propoxyphene as a result of the intervention (Soumerai et al, 1987a). A further study from Northern Ireland reported a downward trend in barbiturate use in line with the secular trend that had existed before the campaign (King et al, 1980). It seemed the campaigns were most successful when the medications had serious adverse reactions and were unlikely to have potential for abuse. Scientific media support was also important (Soumerai et al, 1989).

2.3.5. Mass media campaigns

A Cochrane review assessed the impacts of mass media campaigning on the utilisation of health services (Grilli et al, 2002). It included 20 studies and concluded media campaigning could result in intended influence on utilisation. The authors argued that in order to achieve the objectives, efforts should be

made to ensure that media were used appropriately. The review identified few studies with acceptable quality attempting to evaluate the effects of mass media campaigning on prescribing.

The authors included two ITS studies with prescribing outcomes. A Canadian study evaluated the use of mass media campaigning to reduce calcium-channel blocker prescribing as the first line of treatment for hypertension (Maclure et al, 1998). Another study from the USA assessed the impact of the campaign on reducing the use of aspirin in children to prevent Reye's syndrome (Soumerai et al, 1992). Both studies demonstrated significant reductions (Grilli et al, 2002). The reviewers noticed that most campaigns in areas other than prescribing were meant to increase health care utilisation.

2.3.6. Participatory guideline development

There was very limited evidence on the effectiveness of this intervention. A well designed British BIB study of guideline implementation for five common paediatric conditions concluded that participatory guideline development improved prescribing for all five conditions by 8% on average. The groups receiving the guidelines, but not involved in its production, did not improve as such (North of England Study of Standards and Performance in General Practice, 1992a; North of England Study of Standards and Performance in General Practice, 1992b). This study was included in two systematic reviews (Grimshaw et al, 1995; Wensing et al, 1998). In another British study a group of GPs developed audit criteria for monitoring patients receiving digoxin. They distributed the criteria and discussed them with other GPs. One year follow up suggested that only those who were involved in the development of the criteria actually implemented it (Anderson et al, 1988). This was included in Grimshaw et al (1995). Another review, Soumerai et al (1989), included a USA CBA study. The study observed that an intervention involving participatory guideline development improved prescribing in the development site in comparison to the control, but did not improve prescribing in another site that received the guideline. Also the observed effects in the development site were mainly due to

the adherence of those physicians who were directly involved in the development of the guideline (Bush et al, 1979). The findings suggested that on the whole the intervention might not be cost-effective, if only those who developed a guideline actually used it. One explanation for this questionable effectiveness was that participatory guidelines might not be seen as credible by non-participating clinicians (Grimshaw et al, 1995).

2.3.7. Inter-professional shared care and inter-professional relationship

A Cochrane review assessed the effects of on-site mental health workers (including clinical psychologists) in primary care on the care provided for patients and provider performance (Bower and Sibbald, 2000; Bower and Sibbald, 1999). The review assessed the effects of the intervention on the care provided by physicians for the patients directly receiving shared care ('direct effects') and for those patients who were not subject to shared care ('indirect effects'). They also identified two models of on-site mental health worker care: the substitution model and the consultation-liaison model, although there were some degrees of overlap between the models. In total, 38 studies were included in the review of which 28 studies were considered eligible for prescribing outcome analysis (Bower and Sibbald, 1999). The substitution model resulted in some improvement in direct care, but no change in PCP prescribing was observed in three CBA studies assessing indirect care (Pharoah, 1996; Coe et al, 1996; Baker et al, 1996). The reviewers concluded that there was evidence of no effect from the substitution model of mental health care intervention on PCP prescribing (Bower and Sibbald, 1999). For the consultation-liaison model the evidence was mixed, both for 'direct' and 'indirect' effects on patient care (see Appendix II-2). The reviewers concluded that (modest) effects were likely to be observed if the mental health worker liaison-consultation model was part of a multi-faceted intervention.

The effects of the interventions to improve inter-professional relationships on PCP behaviour are not clear. A Cochrane review of doctor-

nurse relationship included two primary studies, neither conducted in primary care settings (Zwarenstein and Bryant, 2000). The evidence from secondary care suggested improved doctor-nurse relationship might enhance process and patients outcomes. Lack of evidence in this area might be due to the complexity of design and evaluation of the intervention. Inter-professional relationships and interactions are essential parts of modern primary care and more evidence in this area is required.

2.3.8. Educational outreach visits

A Cochrane review included eighteen RCTs of educational outreach visits or academic detailing (Thomson O'Brien et al, 1997). Thirteen studies were on prescribing. Positive effects were observed from all prescribing studies (some non-significant). The reviewers identified four different models of educational outreach visits. The review concluded that educational outreach visits were effective in reducing inappropriate prescribing. The effects were usually small to moderate but of clinical importance. The cost-effectiveness of the intervention, especially models involving outreach visits plus other interventions or those with subsequent visits was not well established. Nine included studies had prescribing outcomes and were conducted in primary care. A Belgian study of GPs evaluated the effect of one educational visit by a specially trained GP and educational materials compared to no intervention to reduce benzodiazepine prescribing (Berings et al, 1994). A Swedish study attempted to improve prescribing of lipid lowering drugs for patients with hyperlipidaemia (Diwan et al, 1995). A UK study assessed the effectiveness of a single outreach visit and educational materials compared to no intervention in improving NSAIDs prescribing. The positive effects lasted for at least five months (Newton-Syms et al, 1992). Another British study used educational outreach visits as well as clinical guidelines (educational material) and follow up prompts to improve quality of asthma care. The intervention resulted in improved quality of asthma prescribing (Feder et al, 1995). A USA study reported reduction in the use of propoxyphene, cerebral and peripheral vasodilators and cephalexin in physicians identified as high prescribers (Avorn and Soumerai, 1983). The

improvements lasted for at least nine months and were cost-effective (Soumerai and Avorn, 1986). A study conducted in Indonesia compared two interventions (educational outreach visits and formal seminar, both accompanied by educational material) with no intervention to improve drug use in the management of acute diarrhoea in children. They reported that the seminar resulted in significantly greater changes than the educational outreach visits. Both interventions resulted in significant improvement compared with the control. The outreach visits were less costly than the seminar ($0.77 US versus $3.30 US per participant) (Santoso et al, 1996). One trial of outreach visits plus audit and feedback reported important improvements in prescribing (McConnell et al, 1982). A USA study attempted to reduce the use of high cost NSAIDs and increase the use of ibuprofen and salicylates by providing clinical pharmacy services in an HMO. The intervention was not cost-effective (Stergachis et al, 1987). An Australian RCT of GPs aimed to reduce benzodiazepine prescribing. The relative difference between the intervention and control groups was not statistically significant (Yeo et al, 1994; de Burgh et al, 1995).

The above systematic review included RCTs only. Some of the trials were also included in other reviews (Chrischilles and Gondek, 1997; Le Grand et al, 1999). A more recent review (Gross and Pujat, 2001) included another RCT from Australia that had found educational outreach increased the use of effective and cheap antibiotics and reduced the use of more expensive antibiotics (Ilett et al, 2000). Other reviews provided evidence of effectiveness from other designs. Wensing and Grol (1994), as well as others, included two papers based on a CBA study that found educational outreach an effective intervention in reducing long term diazepam prescribing (Ray et al, 1986) and inappropriate antibiotic prescribing (Schaffner et al, 1983). Another review (Grimshaw et al, 1995) identified a CBA study using educational outreach to implement a clinical guideline for improving anti-ulcer treatment. Checking the original paper revealed that statistically significant improvement was observed one month after the intervention. However, the improvement was not anymore significant two months after the intervention (Raisch et al, 1990). A meta-analysis of interventions to change prescribing concluded that one-to-one meetings (educational outreach) were effective methods of improving prescribing (Pippalla et al, 1995).

2.3.9. Audit and feedback (with or without peer review)

A Cochrane review assessed the impact of audit and feedback on provider performance and patient outcomes (Jamtvedt et al, 2003). The review included 85 RCTs. The included studies covered a range of different audit and feedback interventions including performance feedback, cost feedback, written or verbal feedback and peer review feedback. The review's overall conclusion was that audit and feedback might be effective, but the supportive evidence was patchy and the effects tended to be small to moderate. It also concluded that the effectiveness was correlated with baseline compliance (i.e. the lower the compliance, the more effective the intervention) and had no correlation with the complexity of the target behaviour. Among all studies included in the review seven studies were conducted in primary care and had prescribing outcomes. Two studies compared 'audit and feedback alone' with 'no intervention' (O'Connell et al, 1999; Mainous et al, 2000). Both studied observed no significant effects from the intervention. Audit and feedback plus educational outreach were effective in reducing tetracycline prescribing (McConnell et al, 1982) and similar interventions were effective in a multi-centre study in three (out of four) European countries to improve prescribing for asthma (Veninga et al, 1999). One multi-faceted intervention was ineffective in reducing the average number of prescribed drugs (Kafuko et al, 1999). Canadian researchers evaluated audit and feedback plus educational material intervention against no intervention control. As a result, reduction in prescribing costs and improvement in the choice of first line drug was achieved in the intervention group (Hux et al, 1999). This RCT was also included in Gross and Pujat (1999) review. In a separate review, Lomas and Haynes concluded that computerised audit and feedback mechanisms improved prescribing in primary care (Lomas and Haynes, 1988).

Other reviews included non-RCT evidence. Wensing and Grol (1994) identified two CBA studies assessing the effectiveness of peer review audit and feedback. One study observed that the intervention group enjoyed less increase in prescribing charges than the control (Harris et al, 1985). The second study

observed a decrease in prescribing costs per patient per DDD as a result of the intervention, while there was no significant decrease in the amount of prescriptions (Lassen and Kristense, 1992). A later update of their systematic review concluded that single interventions using feedback were more effective than feedback plus information transfer (Wensing et al, 1998). The review by Mugford et al (1991) included a British CBA study assessing the impact of meetings plus number and cost of prescription feedback to GPs twice a year for two years. The intervention group's prescription rate per patient fell more than the control group (Harris et al, 1984). Gill et al (1999) found audit and feedback an effective method of behaviour change. Beilby and Silagy (1997) reviewed the effects of costing information feedback on performance. Their review did not provide any further evidence of effect than what already discussed.

2.3.10. Reminders (including computerised systems)

A systematic review of computerised advice on drug dosage included fifteen studies all conducted in secondary care (Walton et al, 1999; Walton et al, 2001). The reviewers failed in identifying any studies conducted in primary care. The reviewers concluded that computer support improved achievement of target drug dosages with less adverse outcomes.

Most evidence on the effectiveness of reminders originated from preventive care (Hunt et al, 1998; Mitchell and Sullivan, 2001) and there was limited evidence for prescribing. Soumerai et al (1989) included an ITS study that followed the implementation of a reminder system for the antibiotic treatment of streptococcal pharyngitis. Four years of time series data demonstrated marked reduction in the percentage of untreated patients as a result of computer generated reminders. The effect disappeared as soon as the reminders stopped (Barnett et al, 1978). A systematic review assessed the impact of computers on primary care consultations (Mitchell and Sullivan, 2001). The review included two interventional studies with positive effects from computerised reminders on prescribing (McDonald et al, 1980; Jones et al, 1996), but the evidence was not strong (see Appendix II-2). Using the evidence

from preventive care and prescribing in secondary care, it could be extrapolated that the reminder systems may be effective in specific areas of prescribing, for example targeting poly-pharmacy, preventive prescribing or drug interactions where it is easy for the physician to forget, and a simple reminder can be of great help. In these scenarios reminders would correct the errors of omission via acting as secretarial reminders (Soumerai et al, 1989).

2.3.11. Patient mediated interventions

A Cochrane review assessed the effectiveness of the interventions to promote patient-centred health care in improving patient-provider consultation process, provider performance and patient outcomes (Lewin et al, 2001). Seventeen studies were included in the review. One RCT was performed in primary care and reported prescribing outcomes (Clark et al, 1998). In this study patient-centred training for providers plus condition-specific training for both providers and asthmatic patients and their parents was evaluated against no intervention. Parents of children in the intervention group were significantly more likely to report that the paediatrician had prescribed inhaled anti-inflammatory medicine for the child and also that the paediatricians had given the family a written plan for adjusting the treatment at home when symptoms change. Gill et al (1999) included four studies involving patient mediated interventions to improve primary care prescribing. All reported positive outcomes. A further Cochrane review of audit and feedback (Jamtvedt et al, 2003) included an RCT that compared audit and feedback alone with audit and feedback plus patient education and concluded adding patient education did not change the outcome (Mainous et al, 2000).

2.3.12. Local opinion leaders

A Cochrane review assessed the impacts of interventions using local opinion leaders on provider performance and patient outcomes (Thomson O'Brien et al, 1999). The reviewers identified and included eight high quality primary studies.

None of the studies were conducted in primary care. Seven out of eight studies used a sociometric questionnaire with questions on humanity, communications and knowledge for identification of opinion leaders (Hiss et al, 1978). The review identified mixed effects as a result of the intervention. It also noticed that the activities of opinion leaders were not always known; so even if there was evidence of effect it would be difficult to replicate the intervention elsewhere. The absence of any high quality study from primary care may be the result of the inherent difficulties in identifying opinion leaders in primary care (Thomson O'Brien et al, 1999). An earlier systematic review had also failed to identify any evidence for the effectiveness of local opinion leaders or 'educationally influentials' in primary care (Lomas and Haynes, 1988).

2.3.13. Continuous quality improvement (CQI)

A systematic review assessed the effects of CQI programmes on provider performance and patient outcomes in different health care settings (Shortell et al, 1998). The investigators included 42 single-site primary studies of which only two were randomised trials. They also included 13 multi-site studies of which one was RCT, one was CBA and three were still in progress. Only one study met the overview's inclusion criteria. A three-arm RCT in the USA compared educational visits plus CQI with educational visits alone (Goldberg et al, 1998). The control group received no interventions. The outcomes were prescribing for hypertension and depression. The arm including CQI did not improve over and above the other intervention group. The review concluded that the findings of primary studies could not be linked to the CQI because of weak designs. It argued that the observed failure of the CQI in improving the services were due to inappropriate choice of target behaviour or poor implementation of CQI (Shortell et al, 1998). Both arguments support the finding that CQI is not appropriate for every scenario and may be ineffective.

2.3.14. Financial incentives

Four systematic reviews assessed the impacts of different forms of financial incentives on physician behaviour (Giuffrida et al, 1999; Gosden et al, 2000; Chaix-Couturier et al, 2000; Gosden et al, 2001; Armour et al, 2001). A systematic review of different financial incentives included 89 primary studies of which seven were RCTs (Chaix-Couturier et al, 2000). Only one RCT (conducted in the USA) was on primary care prescribing, in which the authors assessed the impact of disclosing the threshold for financial sanctions on provider behaviour (Nyman et al, 1995). The disclosure resulted in reduction in prescriptions. The reviewers also used evidence from other studies and warned that financial incentives should be used carefully as they might cause conflicts of interests between different providers and also between the health system and other public sectors (Chaix-Couturier et al, 2000). The second review (Armour et al, 2001) assessed the effects of bonus systems and included seven studies of which one was ITS and two were RCTs. None were on prescribing in primary care. The authors concluded that bonus payments and explicit financial incentive had mixed effects on provider behaviour and on quality of care.

Two further reviews were published as Cochrane reviews (Giuffrida et al, 1999; Gosden et al, 2000) and were also jointly published elsewhere (Gosden et al, 2001). One review assessed the impact of target payment on PCP behaviour and managed to include only two studies (both on immunisation); only one of them demonstrated significant change as a result of target payment (Giuffrida et al, 1999). The authors identified no studies of prescribing. The other systematic review assessed the impact of different remuneration systems (fee-for-service, capitation and salary) on PCP behaviour (Gosden et al, 2000). It included four primary studies and concluded that fee-for-service was likely to increase utilisation. The conclusion was mainly based on those studies (three out of four) that had no prescribing outcomes. The only study with prescribing outcome was a large CBA study from Denmark assessing the effects of adding fee-for-service payment to capitation on repeat prescription rates (Krasnik et al, 1990). Contrary to the investigators' expectations, statistically significant reduction in repeat prescription was observed in the intervention group while no such effect was observed in the control group. The reviewers interpreted this

finding as implying '*PCPs do not respond to financial incentives or the level of the fee was not sufficient to encourage such behaviour*' (Gosden et al, 2000, p 8). Neither explanation easily justified the significant *reduction* that followed the introduction of the fee.

In general four systematic reviews failed to provide solid evidence on the effects of different financial incentives on primary care prescribing. Given that it is possible to design good quality studies to evaluate the effects remuneration systems on prescribing, it is surprising that the research evidence from high quality studies is so scarce.

2.3.15. National drug policies, essential drug programmes and regulatory approaches

One systematic review of national drug policies and essential drug programmes failed to identify any study from developing countries that satisfied the minimum requirements of this overview in terms of the research design (Ratanawijitrasin et al, 2001). However, the reviewers' concluded that essential drug programmes in developing countries might result in more appropriate use of the medications in primary care. It was worth noting that these programmes in developing countries usually involved multi-faceted interventions including increase in the supply of essential drugs as well as education and training. Ratanawijitrasin and colleagues also assessed the effects of policies such as de-registration, upward reclassification (e.g. from OTC to prescription only) or downward reclassification (e.g. from physician prescription only, to other health professionals' prescription or OTC). Again no reliable evidence was identified, but it hinted that unintended outcomes were likely. Unintended outcomes included inappropriate use of other drugs in response to de-registration or upward reclassification. The reviewers also found that de-registration on its own was not equivalent to the removal of a medication from use. Unintended outcomes of downward reclassification included increased dispensing of the drugs by unqualified personnel (Ratanawijitrasin et al, 2001). Although the above mentioned consequences seemed logical, they were not evaluated in properly designed studies. This was

especially required in order to quantify cost, benefits and consequences of these policies. For example downward reclassification of oral re-hydration solutions (ORS) for the treatment of diarrhoea in many countries might have resulted in inappropriate use of the products, but the benefits might have outweighed the adverse consequences.

Another group of interventions involved restricting physician prescribing choice by limited formularies. This method could be the result of national policies or take place locally (e.g. in PCOs). A meta-analysis of prescribing interventions concluded that prescribing restrictions by formularies were effective (Pippalla et al, 1995). This finding was based on three primary studies that Pippalla and colleagues included in their review. Unfortunately, the methods and findings of the review were not fully reported. It was not even possible to establish which primary studies had been used.

2.3.16. Multi-faceted interventions

A systematic review of prescribing by Gill et al (1999) concluded that multi-faceted interventions were successful in just less than half (49%) of the interventions. Similarly, analysis of studies from primary care setting would have concluded that only 10 out of 19 interventions resulted in intended improvements in prescribing (Gill et al, 1999, Table 3). This achievement rate was less than 'patient mediated interventions' and 'audit and feedback' and was comparable to 'mailed printed educational material only' interventions (Appendix II-2). Gross and Pujat (2001) concluded that multi-faceted interventions were effective methods of improving antibiotic prescribing. They included three CBA studies using multi-faceted strategies, all of which reported improvements in prescribing (Perez-Cuevas et al, 1996; Gonzales et al, 1999; Saint et al, 1999).

A systematic review concluded that feedback only interventions were more successful than interventions that comprised feedback and information transfer (Wensing et al, 1998). In line with this conclusion a Cochrane review of audit and feedback used the evidence from a few studies to conclude that there

was no evidence suggesting that multi-faceted interventions were more effective than audit and feedback alone (Jamtvedt et al, 2003). Similar conclusion could have been made based on evidence from primary care prescribing studies only. An RCT in primary care compared CQI plus educational visits with educational visits and assessed prescribing for hypertension and depression (Goldberg et al, 1998). The CQI plus educational visits had no advantage over educational visits for neither of the outcomes (Shortell et al, 1998). Angunawela et al (1991) reported the results of a cluster randomised trial in which the newsletter only intervention was as effective as the newsletter plus seminar intervention in reducing antibiotic prescribing (-7.4% versus -7.3%).

Another Cochrane review of on-site mental health worker care effects on PCP performance suggested that the interventions were more likely to succeed in changing antidepressant prescribing if the intervention was related to multi-faceted consultation-liaison interventions involving education, patient-based consultation and feedback and reorganisation of services (Bower and Sibbald, 1999). The effects of these interventions tended to be limited to patients directly in the care of mental health care workers; PCPs did not seem to change their behaviour when caring for other patients.

2.4. Discussion

2.4.1. Limitations

The search strategy might have been improved by explicit searches of other databases. Limiting the main searches to certain databases might have resulted in loss of relevant reviews. However the iterative approaches that followed the original searches, including forward searching and bibliographic searches, could have overcome this limitation. The search strategy deemed reasonably comprehensive since it identified at least two systematic reviews (Pippalla et al, 1995; Chrischilles and Gondek, 1997) that had not been picked up by the earlier overviews of systematic reviews with comprehensive search strategies (NHS CRD, 1999; Grimshaw et al, 2001).

The overview was limited to systematic reviews. Like any other research method, explicit inclusion criteria meant that potentially useful evidence was left out. For example if primary studies were not included in the systematic reviews, they were not identifiable in this method. Given that several systematic reviews were included, little important evidence should have been missed. Despite this, and especially for interventions with limited evidence base, important evidence may still be available from studies that were not included in the reviews. This was particularly true for national drug policies and regulatory interventions (2.3.15). For example a non-systematic review of evidence reported that budgetary restrictions for prescribing in Germany resulted in sharp reductions in the number of prescriptions (Bloor and Freemantle, 1996) hence had positive outcomes. On the other hand, a USA study concluded that policy of payment restriction for prescribed medicines (under Medicaid programme) resulted in under-prescribing of certain drugs, including vital pharmaceuticals such as insulin and cardiovascular medications (Soumerai et al, 1987b). Considering that there was very limited evidence for the effects of regulatory strategies available from the systematic reviews, evidence like this could be informative. Also for educational outreach visits recent studies (Freemantle et al, 2000;

Watson et al, 2001; Freemantle et al, 2002; Banait et al, 2003; Ricordeau et al, 2003; Crotty et al, 2004) that were not included in the overview provided a mixed picture of effectiveness, with some demonstrating no effects because of the intervention. This was different from the systematic reviews' evidence where most included studies found educational outreach effective (2.3.8).

Inclusion of studies in the overview, data extraction and data synthesis were conducted by one reviewer. This could be a threat to the validity of the findings of the overview. Definitions, criteria for inclusion of studies and analytical methods were agreed between the reviewer and his adviser. Attempts were made to reduce the likelihood of bias by re-checking of the included and excluded studies and data extraction forms. Also all the results and analyses were reviewed by the adviser to ensure methodological approaches were followed. The adviser had extensive experience in teaching methods of systematic reviewing and conducting systematic reviews, including reviews of provider behaviour change (Girmshaw and Russell, 1993; Russell et al, 2000). The reviewer received extensive training on the methods of systematic reviewing before the conduct of the overview and since then has taught systematic reviewing techniques and contributed to a few ongoing reviews including three Cochrane reviews.

Although quality of systematic reviews was assessed by attention to the methodological quality of the review and also the included primary studies (Appendix II-2), no attempts were made to assign quality scores to the included systematic reviews. This may have resulted in the loss of understanding of variation between the outcomes of different reviews. It may also have resulted in the loss of understanding of the validity of the reviews' conclusions. Low quality studies tend to report larger effects and quality scoring is a method of examining the validity of the findings (Moher et al, 1998). However, previous studies concluded that formal quality scores of clinical trials of interventions were unlikely to explain the observed variations in effect sizes (Juni et al, 1999; Balk et al, 2002). A recent study examined the relationship between different measures of quality and the treatment effect sizes in RCTs included in published meta-analyses (Balk et al, 2002). The study failed to identify any relationship between the reported effect sizes and the quality measures. The

authors concluded '*it should not be assumed that any given quality measures will necessarily explains the effects found*' (ibid, p 2891). Quality score were less useful in explaining variation in outcomes where the interventions were complicated and less standardised (Balk et al, 2002), which is the case for provider behaviour change interventions. It is not clear how the quality of systematic reviews correlates with the outcomes. However, it is known that even high quality systematic reviews are subject to variations in the interpretation of their findings. In Ezzo et al (2001) study, two independent methodologists interpreted the conclusions of some Cochrane reviews using explicit criteria; then the reviews' authors were asked to provide their interpretation of the conclusions based on same criteria. Ezzo et al demonstrated that the methodologist readers interpreted the reviews' conclusions differently from what the authors had intended. It is also known that high quality systematic reviews do not necessarily reach valid conclusions. A comparison of three high quality systematic reviews of pain relief for patients with sciatica demonstrated important differences between the reviews (Hopayian, 2001). These variations were not captured by Oxman and Guyatt (1988; 1991) quality assessment index. As pointed out by Hopayian, one reason could be that the application of quality assessment tools did not require any knowledge of the included studies. In the overview presented in this chapter due care was given to the primary studies included in the systematic reviews. Also in this overview careful attention was paid to the evidence base of conclusions presented in the systematic reviews. This should have overcome some of the concerns about understanding the quality of included systematic reviews. The consideration of primary studies included in the reviews was conducted also because of technical reasons. The overview focused on physician prescribing in primary care, while many reviews had wider inclusion criteria. Therefore, the results of even high quality systematic reviews could have been invalid in here if they were based on studies other than PCP prescribing.

Some of the systematic reviews were already out of date in 2001 and still were not updated in 2004. This was particularly striking for the Cochrane reviews. It seemed that the Cochrane review groups were facing difficulties in

achieving the objective of providing timely and valid systematic reviews of evidence in the area of health care provider behaviour change.

2.4.2. Interpretation of evidence in the systematic reviews

The evidence from primary investigations was not consistently interpreted in different systematic reviews. This was perhaps the most important concern for the validity of the findings of the included systematic reviews. A very interesting example was observed from a study (De Santis et al, 1994) included in three systematic reviews. This was an Australian RCT of mailed educational materials followed by educational outreach visit to improve antibiotic prescribing for tonsillitis in primary care. One systematic review referred to this as a study which observed little improvement as a result of the intervention (Gross and Pujat, 2001). Another systematic review used a much more positive tone and reported that the intervention group resulted in better compliance with targeted behaviour than the control group (Anderson and Lexchin, 1996). A further review again reported the findings as positive, but made the mistake of identifying the review as a CBA study rather than RCT (Chrischilles and Gondek, 1997).

A further example was based on the interpretation of the findings of an American CBA study in different systematic reviews (Schaffner et al, 1983). The study had four arms: educational outreach by pharmacist, educational outreach by physician, mailed printed educational material and no intervention control group. A systematic review of antibiotic prescribing referred to this study as a study that demonstrated both pharmacist-led and physician-led educational outreach as effective programmes (both statistically significant compared with the control) although the physician-led approach was more effective (Gross and Pujat, 2001). An earlier systematic review of ambulatory care prescribing referred to this primary study as if it demonstrated the pharmacist-led educational outreach was not effective, and then tried to dispute the finding using other information (Soumerai et al, 1989). Both reviews had acceptable quality, but paid attention to different details from the same study. Another

primary study (Ray et al, 1986) was a matter of dispute between systematic reviewers. Three systematic reviews emphasised the positive finding of the study: reduction in diazepam prescribing for long term users in the intervention group (Wensing and Grol, 1994; Grimshaw et al, 1995; Yano et al, 1995). Conversely another systematic review paid attention to the finding that overall diazepam prescribing in the intervention and control groups remained the same despite the intervention (Soumerai et al, 1989).

Another study was included in two systematic reviews (Gehlbach et al, 1984). This primary study was conducted on residents and therefore its evidence was beyond the scope of the overview. It was however referred to in this overview since the variation in interpretation provides an important lesson. A *BMJ* systematic review (Mitchell and Sullivan, 2001) concluded from it that computerised feedback increased generic prescribing. Details of feedback messages were not reported in the review or in the additional table provided on the *BMJ* website. On the other hand, Beilby and Silagy's (1997) review concluded from the same paper that the provision of cost information computerised feedbacks to GPs would increase generic prescribing. It seemed that Mitchell and Sullivan (2001) had taken the liberty of choosing what to report and what not to report from the primary data. This incomplete reporting was in line with the objectives of their review: the effects of computers on behaviour. However, their incomplete reporting deprived readers from understanding potential interactions between the content of the feedback and its computerised delivery.

The flaws in the results went beyond the interpretation of the primary studies. There were occasions when the systematic reviews got to conclusions that were based on unsafe interpretation of the results and the reviews had provided no data in support of the claim. For example Beilby and Silagy (1997) concluded that '*academic detailing may be more appropriate for prescribing and computerised feedback for test ordering*' (p 92). To get to this conclusion, head to head comparisons of computerised feedback and academic detailing for *both* prescribing and test ordering were required. Beliby and Silagy did not report any such comparisons. Indeed they reported two trials (one on computerised

feedback and the other on academic detailing) that both succeeded in improving prescribing behaviour (Appendix II-2).

The systematic review by Gosden et al (2001) suggested the effects of remuneration systems on prescribing were small and not as economic theories suggested. In contrast, an earlier review concluded that financial incentives were effective interventions for changing prescribing in primary care (Chaix-Couturier et al, 2000). The main reason for these conflicting conclusions was that the reviewers used different methods. While the former review excluded studies other than RCT, CCT, CBA and ITS, the latter review included other study designs that were more prone to bias (Appendix II-2).

2.4.3. The effectiveness of multi-faceted interventions

The claim that multi-faceted interventions are more effective than single interventions is not new (Lomas and Haynes, 1988) and is probably correct in general. Most previous overviews concluded that multi-faceted interventions were the most effective approaches in changing behaviour (Bero et al, 1998; NHS Centre for Reviews and Dissemination, 1999; Smith, 2000; Grimshaw et al, 2001). The logic is that if a specific behaviour is targeted from different fronts and the factors that hinder the behaviour are tackled using different strategies, then it is more likely to achieve the behaviour change. However, there are a few important issues that should be taken into account when considering the effectiveness of multi-faceted interventions.

First, multi-faceted interventions might not be as cost-effective of single interventions. By definition the cost of implementing a multi-faceted intervention including interventions A and B is more than implementing either A or B. Therefore, if the incremental benefits of adding B to A is less than the extra costs involved, then it may be wiser to stick to the single strategy and use resources for more efficient interventions.

Second, the overview findings did not support the claim that '*the degree of change appears to be directly proportionate to the number, type, and intensity of interventions proposed*' (quoted from Davis in Beaudry 1989, p 286). The evidence suggested that *some* combinations of interventions were more effective than single strategies (Wensing and Grol, 1994; Wensing et al, 1998). But the evidence did not suggest that the relationship was linear. On the contrary, there was evidence suggesting that some combinations did not do better than single strategies (Wensing et al, 1998; Jamtvedt et al, 2003). Also there was no linear relationship between the intensity of an 'effective' intervention and its incremental effectiveness. In a rare attempt to quantify the link between the intensity of behaviour change programmes with their effectiveness, a systematic review (excluded from this overview since it was updated by other reviews) compared the effectiveness of CME programmes of different lengths of from one day to more than a year (Beaudry, 1989). The reviewer concluded that there was a link between the intensity of CME programmes (in terms of length) and their effectiveness. However, a close look at his analysis suggested that the only true difference between the length of programmes and effectiveness was between programmes of one day length and others with any length of time. The effect sizes of CME programmes lasting from 2-6 days to more than a year were very close to each other with large standard deviations (ibid, p 295).

Third, one should not forget that the absolute costs of implementation matter to health systems. Out of proportion emphasis on the effectiveness of multi-faceted interventions may discourage decision makers from doing anything at all for a given behaviour, just because they cannot meet the total bill for a comprehensive multi-front strategy. Given that almost in every country resources available for health care are less than public expectation, hugely expensive quality improvement programmes are not welcome. Policy-makers should not be discouraged from looking for less costly (and perhaps less effective) behaviour change strategies.

And finally, the issue of health care system and provider capacity to improve is also important. Not all changes in the process of care translate into improvement in health care. A multi-faceted intervention may be effective when

there is an important gap between the quality of care given to a group of patients and the expected standards. But when quality of care reaches a minimum standard (which should be defined) then doing more of the intervention may waste resources and hamper the good will of clinicians that cannot respond to the initiative as well as before.

On the other hand the evidence (not included in this overview) hints that sometimes very simple and low cost interventions can be effective. A systematic review identified three UK studies (two RCTs and one CBA) in which clinical guidelines for radiological investigations were mailed to GPs without any other implementing activities (Grimshaw et al, 1995). Two of the clinical guidelines were developed locally (non-participatory) and one was a national guideline. All three studies observed significant improvement as a result of the intervention. A further study suggested although the effect size from passive dissemination is usually small, but it may be a cost-effective method of changing physician behaviour (Mason et al, 2001). Also a recent Latin Square design (Pocock, 1983) study compared a multi-faceted intervention with passive implementation of clinical guidelines for asthma and angina in British primary care (Wright et al, 2003). They observed comparable improvements in both arms of the study. These are similar to evidence included in the systematic review about the effectiveness of multi-faceted interventions (2.3.16) and mailed printed educational material (2.3.4).

In conclusion, the frequently repeated message that 'multi-faceted interventions are more effective' can be misleading. It should be replaced with evidence-based messages identifying the effectiveness of specific combinations of interventions in different circumstances. The issue of cost should be highlighted to ensure that meagre resources are not diverted to limited areas of quality improvement.

2.4.4. Definitions and terminology

The terminology and the definitions of widely used terms in the area of provider behaviour change have evolved through the time, making it difficult to introduce

valid and applicable taxonomies of interventions. There was a lot of variation in the interventions for assessing the effectiveness of intervention for changing provider behaviour. Smith (2000) provided several examples of how the definitions might be misleading or difficult to understand. Even when standard terminology was being used, still there was room for variation (Smith, 2000). That was one of the reasons why systematic reviews were difficult to perform, and the overall generalisability of their findings was limited.

> *Undertaking reviews to identify promising implementation techniques is difficult, because of the complexity inherent in the interventions, the variability in the methods used, and the difficulty of generalising study findings across health care settings* (Bero et al, 1998, p 468).

A good example of misleading definitions could be seen in one the most commonly used interventions: continuous medical education (CME). Formal CME programmes were started in 1930s in the US (Beaudry, 1989). They were aimed at increasing the knowledge and competence of participating professionals and hence improving their practices. Since then there had been some changes in the meaning and applicability of the CME. In earlier systematic reviews that were performed since 1980s there was a tendency to include all interventions aimed at improving professional practice under the CME title (Haynes et al, 1984; Beaudry, 1989; Davis et al, 1992; Davis et al, 1995). This inclusive approach was intentional. Davis argued

> *Distinction between the two appear to be artificial, and have the effect of placing CME in its traditional non-practice-related (away from practice) mode, from which position it would appear to be less than optimally effective, or from which it would appear to be designed more for mass consumption than tailored to individual needs* (quoted from Beaudry, 1989, p 286).

Despite this, progress in the field of provider behaviour change resulted in a need for clear and exclusive definitions. Therefore more recent reviews of CME by the same researchers excluded interventions other than formal out of practice educational strategies (Davis et al, 1999). The ambiguity in definition was not limited to the interventions and was present for 'setting', 'provider' and 'behaviour'. For example investigators have assigned different meanings to

'primary care'. Some investigators expanded it to include not only hospital outpatient clinics but also emergency rooms (Soumerai et al, 1989).

2.4.5. An updated taxonomy of the interventions

Different taxonomies of interventions were introduced in the systematic reviews. A frequently cited taxonomy was developed by Lomas and Haynes (1988). In their taxonomy they divided different interventions according to the target (individual; geographic or speciality community) and whether they were service specific or general. Then they categorised the interventions under 'patient-centred', 'educational', 'administrative', and 'economic' subgroups. This taxonomy is more or less out of date. It also has the limitation of putting several of the most frequently used interventions under the category of 'administrative' interventions. Similarly, others categorised interventional strategies for improving drug use into four groups of: educational, managerial (i.e. 'administrative'), financial and regulatory (Le Grand et al, 1999).

EPOC Cochrane Review Group developed a practical and useful taxonomy of interventions. It provided reasonably exclusive definitions for different interventions. The definitions offered by EPOC are provided in many of their systematic reviews (e.g. Jamtvedt et al, 2003; Thomson O'Brien et al, 1997) and they are also presented in the glossary of important terms in this thesis (Glossary). The taxonomy does not provide higher level relationships between different interventions. Another taxonomy of interventions was developed by Grol and Wensing through a series of publications (Grol, 1992; Wensing and Grol, 1994; Wensing et al, 1998). Their taxonomy provided a higher level relationship by dividing the strategies into 'voluntary' and 'non-voluntary'. At the next stage the voluntary strategies were divided into those targeting 'internal motivation' and those targeting 'external motivation' (Wensing and Grol, 1994). At the next level the interventions were categorised as 'competence oriented', 'performance oriented', 'social pressure', 'structural arrangements' and 'financial'. The original taxonomy is several years old now and some of the definitions used in it are not in widespread circulation anymore. For examples they used feedback instead of audit and feedback and individual

instruction instead of educational outreach. Some interventions discussed in this chapter were not included in Grol and Wensing's taxonomy. It had other shortcomings. It was built on the importance of the source of 'motivation' (internal or external) on behaviour change. Although there is theoretical support for the role of motivation in behaviour change, it still lacks empirical backing. The other problem was that some interventions targeted both internal and external motivation (e.g. peer comparison audit and feedback). Despite all these, the structure of the taxonomy was sound and helped distinguishing different features of the interventions. Table 2.1 offers an updated version of the taxonomy, using the evidence from the overview of systematic reviews of behaviour change interventions. It has also been expanded by addition of the interventions not included in the original taxonomy. Along with the taxonomy, the evidence was summarised using a simple coding system in terms of the effectiveness, the costs and the durability of different interventions. The taxonomy presented in Table 2.1 will be further improved using the evidence from the following chapters of the thesis.

The results of the overview confirmed that after about twenty years of systematic research on provider behaviour change there remained a lot of unanswered questions. Summary of the findings including areas of uncertainty in the effectiveness of the change interventions are presented in Table 2.1. Apart from the question of whether interventions are effective in bringing about any changes, there remain more questions on the cost (in its broader sense) of implementing the interventions, and the durability of the changes invoked because of the interventions. Many studies of clinician behaviour have short follow up. Therefore not much is known of the durability of the intervention effects. Occasional studies with long term follow up provide some insight. A three year follow up of a successful intervention for improving GPs' management of depression showed that GP performance was back to pre-intervention levels (Cantillon and Jones, 1999). More accurate estimation of the durability of the effects can also inform decision analysis models and cost-effectiveness analyses. Thus future research should include economic evaluation and longer periods of follow up.

TABLE 2.1. SAGE TAXONOMY OF INTERVENTIONS, VERSION I. STRATEGIES TO IMPROVE PHYSICIAN PERFORMANCE: EFFECTIVENESS, COSTS, AND LIKELIHOOD OF DURABILITY OF EFFECTS FOR PRIMARY CARE PRESCRIBING.

			Intervention	Effect	Cost	Durability
Voluntary	**Internal motivation**	**Competence oriented**	CME (didactic)	0	Low	Short
			CME (interactive)	++	Medium	?
			IP education	?	Medium	?
			Mailed printed material	+/0	Low	?Short
			Mailed national warnings [$]	+	Low	Medium to long
			Participatory guideline devel.	+	High	?Medium to long
			IP shared care (substitution)	0	?High	Short
			IP shared care (consult.-liaison)	+/0	?High	?
			Mass media	+	Low	?Medium
	Performance oriented		Audit and feedback[$]	+	Low to medium	?
			Reminders (usually computer sys.)	?+	Low to medium	Short
			Educational outreach	?++	Low to high	Short to long
	External motivation	**Social influence**	Peer review	?	?Medium to high	?
			Patient mediated	+	Low to medium	?
			Local opinion leaders	?	?Medium to high	?
			CQI [$$]	?	?Medium to high	?
		Physical support	Practice support	?	?	?
			Essential drugs programmes [%]	?+	?	?
		Financial incentives	Financial incentives	?+/-	Low to high	?Short to medium
	Non-voluntary		Reimbursement and budgetary policies	?+/-	?	?
			Rules, obligations	?	?	?
			Restricted formulary	+	?	?Medium to long
			De-registration / reclassification	?+/-	?	?

*Adapted from (Grol, 1992; Wensing and Grol, 1994). IP: Inter-professional, ++: strong evidence suggests positive (intended) effects, +: limited evidence suggests positive (intended) effects, 0: evidence of 'no effect', +/0: variable effectiveness, ?: no evidence of effect, ?+: no evidence of effect, however less reliable evidence suggests positive effects, ?+/-: no evidence of effect/ likelihood of positive and negative (intended and unintended) effects, $ it usually also has an element of social influence, % it usually incorporates competence oriented approaches, $$ it usually has elements of physical support and competence oriented approaches.

The results also question the applicability of unqualified theoretical reasoning to practice before empirical testing. The findings have cautious implications for the effectiveness of some theory based interventions (such as those based on participatory guideline development) and on occasions provides evidence of effects in opposite direction of what suggested by theory (e.g. for financial incentives). On the other hand it demonstrated that other theory based interventions that had been under scrutiny for longer periods the evidence was in line with theoretical expectations. Interestingly didactic CME was almost always ineffective in changing behaviour, while interactive CME was effective. For many others there was simply no evidence of effect for PCP prescribing (for example local opinion leaders, peer review, CQI). It is just repeating the already known reality that human behaviour is complex and theoretical reasoning may need fine tuning before it is applied to real practice. Unfortunately the ability to perform this fine tuning is very limited since most of research on quality improvement and provider behaviour change is conducted without attention to the theoretical bases.

The results once again demonstrated the level of variation in the effectiveness of the interventions. Many interventions vary in their effectiveness in different contexts and settings. A selective review of theories of behaviour change is presented in Chapter 3 to enhance the understanding of behaviour change interventions and to inform the following stages of the study. Qualitative research methods are powerful strategies in identifying the reasons for variations, especially as it is not possible to devise interventional studies to assess every possible source of variation. Chapter 4 reports the findings of a qualitative study aimed to explain why some clinical guidelines and pieces of evidence appear to have been more successfully implemented in general practice and changed prescribing. The findings of Chapters 3 and 4 are also used for improving the taxonomy on interventions presented in 2.1.

Theory, research, and practice are a continuum along which the skilled professional should move with ease. ... Theory and research are not solely the province of the academic, just as practice is not solely the field of the practitioner.

Glanz, Rimer and Lewis, 2002a, pp 22-23

A theory has only the alternative of being right or wrong. A model has a third possibility: it might be right but irrelevant.

Manfred Eigen; quoted from Diamond, 1989, p 253.

Chapter 3: Theories and models of behaviour change

3.1. Definitions and objectives

3.1.1. Why this review of important theories was required?

The chapter reports a selective review of theories and models to facilitate understanding of guideline implementation and prescribing behaviour change. Specifically those theories were reviewed that could enhance understanding of the variations observed in the effectiveness of different interventions and the results of the overview of systematic reviews presented in previous chapter (Table 2.1). The links between the theories and interventions were highlighted in the text after introducing each theory (or group of relevant theories). The links were then summarised in Table 3.3 presented towards the end of the chapter. The chapter ended with summarising why theory of planned behaviour (TPB) was chosen for detailed assessment in the following stages of the project.

An extensive literature search was unable to find an authoritative review article or textbook which outlined major theories that explained variations in clinician behaviour. The most useful textbook identified was Glanz et al (2002)

on health promotion and education. Although some important theoretical approaches, especially economic theories were missing from this text, a range of psychological, sociological and organisational approaches were presented and compared. Among the review articles, Grol's review of theoretical approaches for practice change was very useful to start from, but it was far from complete and lacked details (Grol, 1997). Ferlie (1997) covered only the theories applicable to the level of organisation and above. NHS CRD (1999) included brief overview of a few theories of behaviour change at individual and organisation levels. Moulding et al (1999) considered handful of theories in developing their model for clinical guideline implementation. Ashford et al (1999) wrote a thoughtful, but less recognised, review. It covered relevant theories briefly. The paper left out some important and increasingly used theories such as the TPB.

3.1.2. What is 'theory' and why theories are required?

Defining the term 'theory' is not straightforward. It has several definitions which vary between and within academic disciplines. Some famous and more recent definitions of the theory are presented in the Table 3.1. Within the definitions offered in the table, the first definition is probably the most objective one since it sets criteria for a theory.

TABLE 3.1. DEFINITIONS OF THEORY (ADAPTED FROM GLANZ ET AL, 2002B).

Definition	Source
A set of interrelated constructs (concepts), definitions, and propositions that presents a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting phenomena.	Kerlinger, 1986, p 9
A systematic explanation for the observed facts and laws that relate to a particular aspect of life.	Babbie, 1989, p 46
A set of relatively abstract and general statements that collectively purport to explain some aspects of the empirical world	Chaftez, 1978, p 2

There are more controversies on what 'models' are and how they differ with or relates to theories. Some argue that theories are different since they are more likely to be abstract and take concrete meanings when applied to specific situations (Resnicow et al, 2002). A further distinction is that '*unlike theories,*

models cannot be falsified and pertain more to the context of discovery than to the context of justification' (Davidson, 1998, p 35). Another distinction between model and theory is that the models are likely to incorporate more than one theory into their structure to improve their empirical value. Others tackle the problem of definition by using the term 'mid-range theories', referring to theories which are less abstract and can be applied to practice (Paterson et al, 2001). The two terms are frequently used interchangeably (e.g. see Checkland, 1981; Davidson, 1998; Riemsma et al, 2002). This confusion in the application of the terms is very common (also see a leading theorist's paper: Ajzen, 1998). Hence, the chapter does not separate theories from models.

Health services research is criticised for empiricism. '*While much of health services research is empirical, the more ambitious studies try to avoid the trap of empiricism*' (Ferlie, 1997, p 184). On the other hand there are limitations in theories and their applications. Ajzen once wrote:

> *... our theoretical models have failed to generate many insights that could not have emerged without them. ... these comments are not meant to imply, however, that non-theoretical interventions based on common sense are likely to be as effective as interventions based on existing theoretical models* ... (Ajzen, 1998, pp 738-9).

Much of clinical behaviour change literature is based on the assumption that clinicians change if they are given information (Kanouse and Jacoby, 1988; Marteau et al, 2002). Few studies have tested the theoretical models that acknowledge the psychological and organisational processes which precede provider behaviour (Raisch, 1990a; Grol, 1997; Bero et al, 1998; Marteau et al, 2002). Over-reliance on empirical evidence might result in proposals that offer few new avenues over what is already known (Greco and Eisenberg, 1993). Using theory informed approaches might enhance the effectiveness of clinical guideline implementation in general and appropriate prescribing in particular (Raisch, 1990b; Grol, 1997). Theories give scholars the opportunity of approaching problems even before starting data collection and provide them a framework for thinking (Fuchs, 2000). Despite all these, most of the work for the establishment of theoretical basis for behaviour change has been around what previous research has shown to work (e.g. see Davis and Taylor-Vaisey, 1997).

This seemingly evidence-based approach lacks coherent theoretical basis and therefore may fail to recognise its limitations and weaknesses (Ferlie, 1997).

Individual theories seldom explain all the complexities and delicacies of provider behaviour (Rosko and Broyles, 1988; Davis et al, 1995). This view has been augmented further with the emergence of new concepts such as knowledge transfer (Davis et al, 2003). 'Knowledge transfer' accommodates a variety of theories from different disciplines for the purpose of facilitating provider behaviour change. For health services research there are few things as useful as 'good' theories (Glanz et al, 2002). '*Theory provides a sound basis for action. ... if action is to be effective, the theory must be adequate and appropriate to the task ... a good theory*' (Mullins, 2002a, p 52).

3.1.3. Categorisation of theories

Theories of behaviour change are categorised along different dimensions. The most common approach is to group theories according to their academic disciplines. If this approach is to be followed, then the first task is to identify relevant disciplines. One author argued that 'behavioural science' could be viewed from the perspective of three disciplines of psychology, sociology and anthropology (Mullins, 2002a). He acknowledged that other social sciences and their sub-divisions played their parts and that there was overlap between the disciplines (*ibid*, p 24). To complement the list, theories from disciplines of management, economics, public policy and politics should also be included. Categorisation of the theories along the lines of the disciplines has serious disadvantages. First, it has made it difficult to identify competing theories of behaviour change. Second, it has resulted in discipline specific jargons. For example, Fuchs (2000) claimed that academic disciplines had no shared concepts. Jargons also limit communication within the disciplines as theoretical concepts are assigned different terms in different theories (Smedslund, 2000). And finally, it is difficult to draw the boundaries of different disciplines.

Theories are sometimes categorised according to the level of their application. For example, theories can be divided into three groups of individual level (also known as micro level) theories, group level theories and organisational level theories (meso level) (Iles and Sutherland, 2001; Rashidian, 2004b). In psychology in particular, there is a tendency to add a level between individual and group levels, called interpersonal (Glanz et al, 2002a). Large-scale level (macro level) is also required to accommodate public policy theories (Shiffman et al, 2002). This hierarchical approach has a few shortcomings. One is that many theories are applicable to more than one level. That leads to another disadvantage, which is not all scholars agree on the level of application of the theories. The third shortcoming is that these levels tend to correspond to disciplines, therefore may not add much benefit over the traditional approach.

Ajzen (1998) categorised the theories of human behaviour according to their level of generality. He explained that some theories only applied to specific types of behaviours. He called these models or theories as content-specific. The alternative group of theories applied to range of behaviours and settings and were called content-free. Ajzen viewed TPB and social cognitive theory as content-free theories. There are other methods of classifying the theories in the literature (e.g. see Abraham et al, 1998; Bekker et al, 1999).

In this chapter the theories are introduced in an implicit order of main disciplines. Hence on a few occasions the borders between the disciplines are deliberately ignored, in order to add to the depth of discussion. The review was selective and meant to capture theories more frequently applied to the field of quality improvement via seeking change in clinician behaviour. The chapter starts with introducing theories in the field of organisational behaviour and management. Then relevant theoretical concepts from the economics discipline are introduced. The diffusion of innovation which is a cross discipline theory is explained next. Then sociological and socio-psychological theories are presented. In the end psychological theories are discussed. It should be noted that most theories are originally developed to explain the behaviour of groups other than health care providers, such as consumers, managers and individual health related behaviour. The differences between health care provider behaviour and other types of behaviour should be taken into account.

3.2. Theories of provider behaviour change

3.2.1. Organisational and management theories of behaviour

The study of organisational behaviour has been defined as '*the study and understanding of individual and group behaviour, and patterns of structure in order to help improve organisational performance and effectiveness*' (Mullins, 2002a, p 20). It originates from organisational sociology (Ferlie, 2001) and is '*the systematic study of the behaviour of individuals, groups and organisations*' (Ovretveit, 2001, p 2). Organisational studies have close links with policy analysis. While policy analysis tends to apply to higher (macro) levels of state or government, organisational studies tend to focus on the level of 'firms' (Ferlie, 2001). Organisational approaches focus on change and improvement from the point of view of organisation as a whole. They maintain that change in individual usually follows change in organisation. Therefore, organisational change may be sought for the purpose of facilitating the process of change in individuals, for example to enhance clinical guideline implementation (Curry, 2000). Individuals play roles in change processes, but that is because of their positions within organisations (e.g. as leaders, managers or key people). Change in individual behaviour is also sought for the sake of achieving organisational objectives. Another advantage of the organisational approach is in its ability to theorise sources of pressure on medical profession. These sources of pressure are the triple constraints of new public policies, new governance structures and the desire to remain self-regulated (Sheaff et al, 2003). Some investigators, namely Chester Barnard, tried to develop a new science of organisation connecting organisation theory, economics and law (Williamson, 1995). More recent scholars maintain that although organisational behaviour is closely linked with the general discipline of management, research in this area tends to draw on other disciplines and uses a variety of methods and methodological approaches (Ovretveit, 2001).

Berwick (1996) provided powerful messages on organisational improvements. First he asserted that not all changes in systems were improvements, but all improvements were due to change. His second message was that effective changes (or 'real improvements') came from changing systems and not from changes that happened within systems. System approach towards health care improvement differs with the traditional practices where improvement is sought by tackling the problem areas (staff, instruments, methods). Another development in the field of organisational change is derived from complexity theory (Miller et al, 1998). Health care organisations are collections of independent but interacting systems with unclear or fuzzy boundaries, similar to complex natural systems (e.g. biological or physical; Plsek and Greenhalgh, 2001). Complexity theory puts more emphasis on the link between the different elements of the organisations than those elements themselves (Plsek and Wilson, 2001; Wilson et al, 2001). Hence, emphasis on removing the barriers to change and considering the roles of the different elements within organisation is not as productive as understanding how different elements of complex and adaptive organisations interact and improving those interactions. The theory also suggests that establishing minimum standards (specifications) is more effectives than detailed planning as it provides more room for creative improvement (Plsek and Wilson, 2001).

Another group of approaches originated in organisational and management theories are famously known as 'total quality management' or 'continuous quality improvement' (Berwick, 1989). It has been defined as '*philosophy of continual improvement of the processes associated with providing a good or service that meets or exceeds consumer expectations*' (Shortell et al, 1998, p 594). It is based on four inter-correlated dimensions that are required for its success: strategic, cultural, technical and structural (Shortell et al, 1998). Despite considerable attention and resources attracted by this appealing approach to improving health care systems, its evidence base is still limited (Blumenthal and Epstein, 1996; Shortell et al, 1998; see 2.3.13 and Tables 2.1 and 3.3).

Like some other concepts in the field of management, leadership has been interpreted in different and sometimes confusing approaches (Parker,

1994; Mullins, 2002b) and '*has been the subject of an extraordinary amount of dogmatically stated nonsense*' (Barnard, 1997 reprint, p 89). Barnard (1997) thought leadership was dependant on 'the individual' (i.e. leader), the followers and the conditions (e.g. settings). It is not clear whether the followers are those who follow the leader or those who are expected to follow the leader. What most agree on is that management and leadership, while interrelated, are different. The 7-S model of organisational framework could be used for this distinction (Iles and Sutherland, 2001; Mullins, 2002b). According to the model, managers focus on 'strategy', 'structure' and 'systems'. On the other hand leaders are more likely to focus on soft Ss, which are 'style', 'staff', 'skills' and 'shared goals (values)' (Mullins, 2002b). Ambiguity over the definition makes it a difficult concept for systematic evaluation. The concept of clinical leadership is used for the analysis of the role of opinion leaders in behaviour change and norm transfer (Mittman et al, 1992; see 2.3.12). Theoretically it is possible to assume that improving leadership skills of GPs may improve their influence over other team members. GP managers can also use their leadership skills to promote the values which the health system stands for, and improve the morale of their colleagues (Pendleton and King, 2002). Fulfilling these roles in the current climate of high pressure and lack of resources may prove difficult. Leader influence may also lead to change in practice despite evidence. Therefore, focusing on leadership should always be accompanied by other interventions.

Interest in the concept of leadership (including medical leadership) is growing (Pendleton and King, 2002; Ham, 2003). However, its role in evidence-based behaviour change – apart from what is achieved by opinion leaders – is less clear (Flottorp et al, 1998). In primary care, GPs have traditionally played the leadership role. This role has developed further within the new primary care organisations. GPs' roles as clinical governance leads have been described using the soft bureaucracy concepts (Sheaff et al, 2003). Soft bureaucracy is used for explaining organisations with rigid external appearance and loosely coupled internal practice (as opposed to hierarchical bureaucracy; Williamson, 1995). Sheaf et al (2003) argued that the emergence of primary care groups and trusts (PCOs) has enabled the GPs to play leading roles within soft managerial structures. These roles can be used for the promotion of evidence-

based medicine, although the level of success is uncertain. Soft bureaucracy implies that GPs in managerial roles use three different legitimating strategies to influence their peers while preventing negative reactions (Sheaff et al, 2003). The most common approach is instrumental legitimating, in which the GPs argue the managerial decisions are made to achieve the organisation's broader objectives (e.g. better quality of care). The next legitimating approach is political legitimating meaning that managers' power originates from voluntary transfer of power from subordinates. Soft coercion or liberal legitimating is the approach in which change is promoted on the basis that it prevents the external threats on professional independence (Sheaff et al, 2003). Nonetheless the effectiveness of interventions based on leadership in changing provider behaviour is uncertain (Tables 2.1 and 3.3).

In recent years there has been growing interest in the concept of organisational learning and in attempts to view (or transform) the NHS bodies as (or into) learning organisations (Davies and Nutley, 2000; Carroll and Edmondson, 2002). Organisational learning is defined as '*a process of increasing the capacity for effective organisational action through knowledge and understanding*' (Carroll and Edmondson, 2002, p 51). The appeal of organisational learning in the NHS is in its promise of enabling continuous improvement. With ever changing health technologies and practices, the notion of learning how to learn is bound to be very interesting. Organisational learning is thought to be profoundly dependant on effective leadership (Carroll and Edmondson, 2002). A learning organisation in turn is an organisation that is capable of using different levels of learning to achieve continuous change and improvement, '*an organisation which facilitates the learning of all its members and continuously transforms itself*' (Pedler, Boydell and Burgoyne, quoted from Hicks, 2002, p 358). Five key characteristics are defined for a learning organisation (Iles and Sutherland, 2001): structure (flat managerial hierarchy, team work, systems thinking), information systems (sophisticated), human resource (personnel as creators and users of organisational learning), culture (promote openness, creativity and experimentation) and leadership (effective, open and risk-taking). What is presented makes clear that learning organisation encompasses a variety of complicated theories and concepts hence it is difficult to implement and evaluate. Applying such framework in its full picture to small

organisations (e.g. general practices) requires resources that might not be available to practices but few large partnerships. PCOs may also lack the required level of interconnection with practices to enable them to establish a full learning organisation framework. However, the theory offers some insights that could provide benefits to general practice. Levitt and March (1995) explained that organisations were history dependent and were based more on routines than on intentions. This is an indispensable insight for general practice and is in line with the medical tradition of valuing case reports (March et al, 2003). No evidence of effect was found for the application of organisational learning and learning organisation concepts to improve primary care prescribing (Table 2.1).

Organisational theories helped the development of new ways of re-arranging the workforce in health service by encouraging team working (e.g. see Berwick, 1996; Priestley et al, 2004). Organisational variables are known to influence the implementation of evidence-based innovations (Dobbins et al, 2001). Whether this influence is causal or correlational is less clear. Iles and Sutherland (2001) performed a selective review of organisational change models and theories. They looked for evidence of the effectiveness of more commonly used models and theories and found minimal evidence. This finding is acknowledged by other investigators (Ovretveit, 2001), and similar results were reached in the overview of systematic reviews (Tables 2.1 and 3.3). Better understanding of the nature of organisational change, and identifying effective interventions is a dire need. Attempts should be made to provide evidence base for change management interventions, e.g. from carefully designed case-studies as well as interventional studies (Ferlie et al, 2001; Rashidian, 2004b).

3.2.2. Economic theories of behaviour

The theory of the firm is the economists' theory of organisation and organisational behaviour. This theory has evolved throughout recent decades. According to the neoclassical theorists,

> *the firm is a collection or set of feasible production plans, presided over by a manager who, buying and selling inputs and outputs in a spot*

> *market, chooses the plan that maximises owners' welfare* (Hart, 1995, p 155).

Welfare is usually represented by profit, but in a health care organisation it can be defined as satisfying health needs, cost reduction or achieving quality targets and often combination of these. If a GP principal is considered as the owner and manager of the practice, the model assumes that the GP will aim to maximise his or her welfare by balancing earning and leisure time (Zazove and Klinkman, 1998). The theory of firm does not answer the question of how the firm or organisation maximises its 'welfare', neither does it explain the structure of the firm (Hart, 1995).

The principal-agent theory or agency relationship, developed in late 1970s, introduces a series of conflicts of interest between different players within an organisation which arise from asymmetries of information (Hart, 1995). The principal-agent theory explains that the owners of firm do not have access to the same information that the manager has. Therefore, the owners try to link the manager's profit to their own (e.g. by giving the manager a share of the profit). This theory, like the neoclassical one, is silent about the structure of the organisation. Information asymmetries and their impact on health care markets are widely used in explaining insurer-patient relationship, where three types of information asymmetries exist: adverse selection, hidden information moral hazard and hidden action moral hazard (Jack, 2000). In all of these asymmetries, the problem arises from the unawareness of the insurer from the patient's (insured) previous health status (adverse selection), the patient's status in accordance with its insurance claims (hidden information) and the patient's preventive or dangerous behaviour which may put his or her in less or more need of costly treatments (hidden action). The principal-agent theory has been applied to patient-doctor relationship, where the doctor acts as an agent for '*the often ignorant, insecure and potentially irrational patient*' and the doctor has much more information about the choices available to the patient (Jensen and Mooney, 1990, p 9). The doctor's benefit or preference may not be in line with the patient's best interest or choice, hence an imperfect agency (McPake et al, 2002). Williams (1988) described this imperfect agency relationship as:

> *The PATIENT is there to give the DOCTOR all the information the DOCTOR needs in order that the DOCTOR can make a decision, and the PATIENT should then implement that decision once the DOCTOR has made it.*

The perfect agency exists when 'patient' is replaced by 'doctor' (and vice versa) in the quoted text. A further principal-agent relationship in medical care exists between the doctors and the hospital or primary care managers, or generally the health system (Fuchs, 2000). The health system's best interest may be to see doctors are 'efficient' (Mooney, 1995), but this may not be in line with doctors' preference. There may be other similar relations as doctors are considered accountable to different professional and political bodies (Royal College of General Practitioners, 1985; Pendleton and King, 2002; Checkland et al, 2004). GPs are expected to behave as agency to a minimum of two groups: the patients and the PCOs' clinical and non-clinical managers. GPs also control access to many secondary and tertiary health services through their gate-keeping roles in the NHS. All these issues cause further conflicts as the principals (patients on one hand and managers on the other hand) may not have similar agendas and seek conflicting roles from their agents (i.e. GPs). Some suggest that the gate-keeping role may help GPs to behave more as informed agent on behalf of the patients (Phelps, 2000), but as explained conflicting roles may prevent this from happening. The other potential problem in patient-doctor agency relationship is that GPs may not be aware of all the available treatment, preventive or diagnostic options for the patient (Folland et al, 2004). Clinical guidelines are able to improve this aspect by informing doctors of standard care. Guidelines can also improve patients' awareness of options available to *them* and hence help patients in establishing a more informed relationship with their agents (GPs).

Demand is '*the quantity of a good or service a consumer will purchase at different prices during a given period of time*' (Rosko and Broyles, 1988, p 58). The traditional neoclassical demand theory is based on the assumption that people make their decisions rationally to maximise their utility, based on their preferences and within the boundaries of their income and the prices of goods or services. The price paid for the health or health services can be represented in monetary or non-monetary terms (e.g. waiting time). In this sense each individual (consumer) is a 'rational economic man' or 'homo economicus'. For

these 'rational' individuals, health and health care are only two among many commodities that they have preferences for (Jack, 2000). This rational explanation of demand may not represent health care due to a variety of peculiarities such as imperfect agency relationship. There are other models of explaining demand for health. Grossman's human capital theory asserts that consumers invest in themselves through health, education and training to increase their earning (Folland et al, 2004, p 125). Health is viewed as a 'capital good' or 'stoke variable' in the investment model of demand (Rosko and Broyles, 1988; Jack, 2000). People inherit health and the value of health depreciates over time. Therefore, people seek health care because they seek health. In this model, the demand for health is in pursue of two benefits. One is the benefit in terms of improved utility achieved as a result of better health. The second benefit is called the investment benefit. It refers to the fact that a healthier person spends more time in other beneficial activities such as work and leisure, and better health improves its capacity to benefit from health. This model has become popular at policy level after the World Bank's World Development Report on investing in health (World Bank, 1993; Jack, 2000). In reality, the production of health depends on a range of variables of which only one is health care. Evidence shows that the marginal benefits of health care in the production of health are small, and are '*nearly on the flat of the* [health production function] *curve*' (Folland et al, 2004, p 84).

Demand for health care has a further peculiarity as it can be induced by the supplier (provider), namely the clinician. It was estimated that as much as a quarter of high technology health services provided to patients were not required (Borowitz and Sheldon, 1993). An important fraction of these services might be supplier-induced. Some strikingly huge regional variations observed in medical care were considered to be induced by providers (McPherson et al, 1982; Raskin, 1991; Parchman 1995). Although most studies originate from the USA or UK, the phenomenon is not limited to any single country (Ellison et al, 2003; Rudge et al, 2003; Miranda et al, 2003; Roudot-Thoraval et al, 2003). The outcome of health care is often uncertain and the natural course of a disease without treatment is often unknown for individual patients. These factors and the complexity of medical care result in the analysis of a lot of information by the doctor before a decision is made (Rosko and Broyles, 1988). The way the

physician interprets the information and the way it is presented to the patient has detrimental effect on the demand for health care, hence the term supplier-induced demand. The supplier-induced demand also happens as a result of imperfect agency relationship in which the doctor puts his benefits ahead of the patient's best interest (McPake et al, 2002). Some physicians tend to refer their patients for further investigation or treatment in which the physicians have financial interest (Iglehart, 1991). The rate of this behaviour in an open access health system may be as high as 40% (Pippalla et al, 1995). Theoretically, clinical guidelines can reduce the supplier-induced demand. Clinical guidelines may improve decision making by organising evidence-based information and reduce geographical variation by defining standard care. For example the dissemination of a national clinical guideline resulted in the reduction of surgical operations for the treatment of glue ear in the NHS (Black and Hutching, 2002). On the other hand, demand theory suggests that the observed reduction in surgery for glue ear may have happened as a result of demand shift to other providers, e.g. private sector (Black and Hutching, 2002) or towards other surgical procedures (Borowitz and Sheldon, 1993).

Not all evidence points towards the inducement as the cause of variation in health care. A review of the uptake of different surgical approaches for breast cancer concluded that physician-induced demand was not the main cause of variation (Greer et al, 2002). Organisational variables (e.g. waiting time and expenditure caps) affect the level of inducement independent of physician characteristics (Reinhardt, 1999). Physicians' inducement ability is also limited by factors such as reputation, patient education, other professionals' views, other patients' views and the views of patient's family and friends (Rosko and Broyles, 1988; Greer et al, 2002). Generally speaking, inducement is more likely to happen in fee-for-service health markets.

Demand theory can be applied to the GPs' behaviour of implementing guidelines in another way. A perfect GP can be defined as someone whose aim is to maximise the patients level of health, or their utility, or to maximise the society's level of health or utility (Miettinen et al, 2002; Blank, 2002). All these are legitimate. Assume a (perfect) GP's aim is to maximise patient's health. The patient's health itself is an intermediary to other maximands for the GP such as

professional gain or satisfaction. The maximand may also be quality benchmarks (and accompanying financial rewards) or meeting budgetary targets through reducing further treatment costs. On the other hand, uncertainty implies that GP cannot be 'sure' of the validity of clinical guideline recommendations when applied to the individual patient and of the effects of other alternatives decisions on the patient's outcomes. The outcome of care is also highly dependant on the patient's behaviour. Therefore, even a clear choice of maximand does not necessarily predict the target behaviour or the GP's clinical recommendation. It is also clear that the choice of maximand in health service is not always straightforward. The role of the clinicians and their behaviours will be even less predictable if all the legitimate maximands are considered (Miettinen et al, 2002). It is equally comprehensible to assume many GPs observe no conflict between the patient and societal maximands, if the GPs consider themselves responsible for the health of their patients within certain boundaries especially limited resources (Royal College of General Practitioners, 1985; Blank, 2002).

Health systems may use remuneration strategies to put the patient's outcome maximand in line with the clinician's financial and professional gain, for example by offering bonus payments for better patient outcomes or improvement in process variables (Armour et al, 2001). Economic theories suggest different remuneration systems (e.g. capitation, salary, fee for service, target payment) influence process and patient outcomes depending on the setting and the way the remuneration is applied (Chaix-Couturier et al, 2000; Gosden et al, 2001). Target payment has a long history in British primary care. For example target payment for family planning services started in 1975 and for immunization in 1965. Many believe target payment has been effective in increasing the adoption of some preventive and screening services (Horder et al, 1986). The evidence on the effectiveness of interventions involving change in remuneration systems was not conclusive (see 2.3.14; also Rashidian et al, 2005), although it hinted towards facilitating quality improvement and reducing prescribing costs (Chaix-Couturier et al, 2000). Also, the overview of systematic reviews suggested that the effects of financial incentives on GP behaviour was mixed and sometimes against expectations (2.3.14; Tables 2.1 and 3.3).

3.2.3. Diffusion of innovation theory

New technologies and innovations gradually diffuse among organisations and individuals. It may take several years, even decades, before an innovation is widely used among its potential target groups (Haines and Jones, 1994). Technologies diffuse with different speeds and one technology may be adopted faster in some settings and groups than the others (Berwick, 2003). The diffusion usually follows a logistic S-shaped curve (Rogers, 1995b; Folland et al, 2004). Diffusion of innovation theory describes the process of the dissemination of a new technology through certain channels in a social system (Moulding et al, 1999). It is meant to explain the variations in adoption and to answer questions like who adopts an innovation and how (Folland et al, 2004). The diffusion of innovation theory (Greer, 1977; Greer, 1985; Rogers, 1995a; Rogers, 1995b), focuses on '*the normal processes by which information is received, circulated, and assessed*' (Greer, 1988, p 6). Therefore, the theory is less concerned with *why* some dissemination strategies fail (or succeed) in changing behaviour, than *how* it happens. It suggests that potential adopters go through five stages of knowledge, persuasion, decision, implementation and confirmation to get to the confident use of innovation (Rogers, 1995a; Rogers, 1995b; see 3.2.5 for other examples of stage-based models). The notion of rationality of behaviour change inherent in the diffusion of innovation is one of its major limitations (Moulding et al, 1999). The separation of formed technologies from dynamic ones to some extent rectifies this limitation (Greer, 1988). Greer argued many technologies developed as they diffused, hence called dynamic; and dynamicity affected their adoption. She hypothesised the assumptions of classical diffusion theory would be met for technologies which were fully formed before their release. Formed technologies, e.g. CT scanners and fetal monitors, diffuse fast compared with the bulk of medical innovations which are normally accompanied with a great deal of uncertainties (Greer, 1988).

Economists have offered an extra angle to the theory by adding the profit motive to the existing information channels (Escarce, 1996; Folland et al, 2004). Hence technologies that increase profit would be adopted earlier. Unlike economic markets, most health systems provide little room for doctors to reap the benefits of technologies and doctors have little financial incentive to

innovate (Phelps, 2000). Regulatory systems manage the profitability of new technologies and the monetary and non-monetary costs of their adoption (Folland et al, 2004). Regulatory systems, e.g. NICE in England and Wales, use guidelines to affect the adoption rates and shift the diffusion curves.

Many clinical guidelines are perceived as new technologies. The diffusion theory is used extensively in proposing strategies for the implementation of clinical guidelines (Mittman et al, 1992; Grilli and Lomas, 1994; Rogers, 1995b; Moulding et al, 1999) and the development of other models of research utilisation (Logan and Graham, 1998). Roger theorised that the innovation characteristics influence the diffusion of new innovations. For clinical guidelines, those characteristics are the relative advantages over normal practice, the compatibility of the recommendations with existing beliefs and values, the complexity of the procedure, the 'trialability' of the procedure and the observabality of the outcomes (Rogers, 1995b). Despite Rogers, these characteristics are more than just the characteristics of the innovation and cover the clinical context characteristics too (Kanouse and Jacoby, 1988). 'Trialability' refers to the extent to which a technology can be tested in limited circumstances before the final decision for its adoption is made (Grilli and Lomas, 1994). The complexity of the technology prolongs the time required for confident implementation. The process of gaining confidence in using a new technology is known as the learning curve (Ramsay et al, 2001). Diffusion of innovation suggests lower levels of complexity and higher levels of trialability and observabality would increase the adoption rate. Grilli and Lomas (1994) found positive correlation between the compliance with clinical guidelines and trialability and negative correlation with complexity. Others found no relationship between the complexity of the target behaviour and the effectiveness of the intervention (Jamtvedt et al, 2003; see 2.3.9 and Appendix II-2).

According to the diffusion of innovation, opinion leaders play important roles in the dissemination process (Rogers, 1995a). 'Change agents' (Moulding et al, 1999) or 'idea champions' (Greer, 1988) are very important, but these are not normally opinion leaders. Opinion leaders tend to be conservative in adoption of new technologies, therefore their role in adoption process is prominent when the benefits of the innovation are discernable (Greer, 1988).

Others suggested that opinion leaders may be among early adopters (Conroy and Shannon, 1995), but this is less convincing as early adopters tend to be younger and more risk taking (Folland et al, 2004). The overview of systematic reviews concluded that the role of opinion leaders in changing behaviour was variable and there was no evidence of effect in primary care (2.3.12; Tables 2.1 and 3.3).

3.2.4. Social influence, power and ecology

The social influence theory emphasises the role that important others play in the individual's decision making and behaviour (Moulding et al, 1999; see also 3.2.6). The theory is better understood if social influence and social power are defined. Raven and Rubin (1983) define social influence as any changes invoked in individual's emotions, beliefs, attitudes or behaviours by other individual(s). They also define social power as a '*potential influence, that is, the ability of some influencing agent to affect some target*' (Raven and Rubin, 1983, p 402). The definitions suggest that there is overlap between this theory and theories such as the social cognitive theory and the TPB as they both theorise social influence in similar approach (Bandura, 1986; Ajzen, 1991). The social influence theory has been widely referred to in clinical guideline implementation literature (Mittman et al, 1992; Conroy and Shannon, 1995; Moulding et al, 1999). It suggests interesting avenues for behaviour change. The theory proposes that social influence is usually exercised through different channels of *information*, *reward*, *coercion*, *expertise*, *reference* and *identification* (Raven and Rubin, 1983). The theory also considers the relationship of these different channels of influence with other factors such as the status and credibility of the influencing agent and the ability to monitor the behaviour.

According to the theory, *informational influence* is independent of the influencing agent (Raven and Rubin, 1983). It is plausible to say informational influence is asserted by almost all clinical guidelines as guidelines are information sources. However, the effects of informational influence on behaviour are uncertain and variable (Grad et al, 1997; see also 'mailed printed

education material' in 2.3.4). Unlike information, *coercion* and *reward* are socially dependent on the influencing agent and also on the ability of the agent to monitor the behaviour (Raven and Rubin, 1983). For example, a NICE guideline is more likely to be implemented if NICE is perceived as capable of monitoring as well as rewarding or coercing. This is truer if the adherence to the guideline can be accurately measured. Coercion and indeed reward may take different forms (e.g. soft coercion in 3.2.1, ecological manipulation in here, financial incentives in 3.2.2). Lewin's field theory (force field analysis) is another theory that deals with the effects of coercion on behaviour change (Lewin, 1951). Field theory envisages that increasing pressure may make the desired change more likely to happen. However, opposing forces gradually grow in power and so the behaviour change is inhibited or reversed. The resulting situation will be a stressful environment with limited desired outcomes. Thus, Lewin suggests that to achieve change it is more fruitful to tackle the opposing and inhibiting factors before and instead of increasing the pressure (Lewin, 1951; Raven and Rubin, 1983; Iles and Sutherland, 2001).

Expert, *referent* and *legitimate* influences or powers are dependent on the influencing agent, but not on surveillance (monitoring) (Raven and Rubin, 1983). These themes are close to each other and overlap to some extent. *Referent power* refers to the social influence derived from following suit of peers and behaviour models. If practitioners perceive that their peers follow a clinical guideline, they are more likely to do the same. One example of achieving change through referent power is when anonymous local audit results are circulated among all participating practitioners, while highlighting the difference of the individual practitioners by local norms (e.g. as part of audit and feedback). *Legitimate powers* are observed from those who are seen by other as legitimate sources of request for behaviour change. Concept of soft bureaucracy explained some of the legitimisation approaches that GPs may use (3.2.1). Expertise, or *expert power*, might exert positive or negative influence on behaviour depending on the level of respect for the influencing agent. The influence of consultants on the GPs' prescribing is one example of the expert power (Armstrong et al, 1996). Using these themes, one predicts that clinical guidelines, especially when supported by expert opinion and based on evidence are likely to exert the influence on the physicians' behaviour as they are in the

expert position. On the other hand in many occasions the guidelines may reduce the expertise influence, as they will demystify the basis of the expert opinion and highlight the limitations and weaknesses in the evidence base of recommendations. Overall, the peers' beliefs and norms are considered as a prominent determinant of the individual's behaviour in the social influence (Mittman et al, 1992). Social influence as well suggests that in the presence of uncertainty individuals are more likely to be influenced by their peers (Mittman et al, 1992; see 3.2.2).

Social judgment theory also helps the understanding of physician behaviour in relation to uncertainty (Hammond et al, 1977). Social judgment defines uncertainty as the '*zone of ambiguity ... that is the conceptual space between that which can be observed and that which must be inferred because it cannot be observed*' (*ibid*, p 3). Social judgment suggests that in situations where the decision makers have no direct access to the valid variables that can affect the decision, they use other variables (referred to as 'proximal cues') with limited validity and reliability for making their decisions. Proximal cues are usually obtained from the environment or 'ecological situation' (Dowding and Thompson, 2003). These cues include peers' views and expectations and patients' signs and symptoms. Given that the proximal cues are perceived differently by different individuals (Ajzen, 1991), and also that different individuals are in contact with different cues, variation in practice is likely.

The notion of 'uncertainty' is also very close to the judgmental error. The theories of human inference explain individuals make erroneous social judgments because of 'representativeness', 'availability', 'framing' and 'vividness' (Raisch, 1990b). The first cause of judgmental error is the similarity judgment or 'representativeness', which is when the relationship between the events is misjudged. For example a GP may assign the recovery of the patient from a disease to a certain drug, while the causality is not ascertained. The 'availability' heuristic explains that individuals use readily available memories for the interpretation of new information. 'Framing' explains that the way the information is presented affects the interpretation of it. 'Vividness' suggests that more interesting information is usually perceived as more important or useful.

The understanding of the nature of social judgment can help the methods used to influence prescribing and to explain variations and errors.

Social influence theory recommends that different target groups ('social settings') warrant different strategies. Three types of settings identified in the literature are the interpersonal setting, the persuasion setting (moderate size groups) and the mass media setting (Mittman et al, 1992; Conroy and Shannon, 1995). Mittman et al (1992) proposes that in interpersonal settings, normative and informational social influences through approaches such as academic detailing or apprenticeship are the most effective. In persuasion settings opinion leaders and auditing (continuous quality improvement) are more likely to exert social influence and transfer norms. The strategies used in mass media approaches, e.g. publications in journals, are less likely to be effective, but are the least demanding methods.

Social ecology theory maintains that environmental factors influence the behaviour. In this regard, the theory is similar to many other theories discussed here. However, the theory goes further and adds that individuals also tend to change the environment in an interactive process (Moulding et al, 1999; see also 3.2.5 'social cognitive theory'). One potential usage of this theory is in its application for exerting indirect influence on behaviour through ecological manipulation '*in which one person influences another by altering some aspects of the environment*' (Raven and Rubin, 1983, p G-8). As an example, guideline implementation might be supported by removing a discouraged medicine from the formulary. The main advantage of using social manipulation rests in its ability in changing behaviour without target group noticing. If the target population become aware and are still unconvinced of the need for change, they may put a lot of effort to neutralise the social manipulation, as explained in force field analysis (Iles and Sutherland, 2001). Examples of these negative reactions have been observed in social behaviours such as driving (Raven and Rubin, 1983). In the field of prescribing also the worry is that local or national formulary manipulations may cause problems in which '*the intentions are subverted even if the letter of the law is followed*' (Horder et al, 1986, p 521). Direct regulatory approaches towards prescribing may achieve short term success. However, they neither achieve improvement in future prescribing

decisions, nor solve the route cause of the targeted behaviour (Raisch, 1990b). The social ecology theory further suggests that suitable and concurrent changes in environment variables and individual behaviour will increase the likelihood of successful change (Moulding et al, 1999). In this way, the theory overlaps with organisational approaches.

Any intervention that applies all the concepts put forward by the social influence theory and other theories presented in here inevitably would be multi-faceted. The overview of systematic reviews did not find any evidence of effectiveness of interventions that applied these theories in their totality. Nonetheless, the theories help the understanding of variations observed in effectiveness of several different interventions presented in Tables 2.1 and 3.3, e.g. audit and feedback, peer review, educational outreach, financial incentives and some non-voluntary interventions.

3.2.5. Transtheoretical model and social cognitive theory

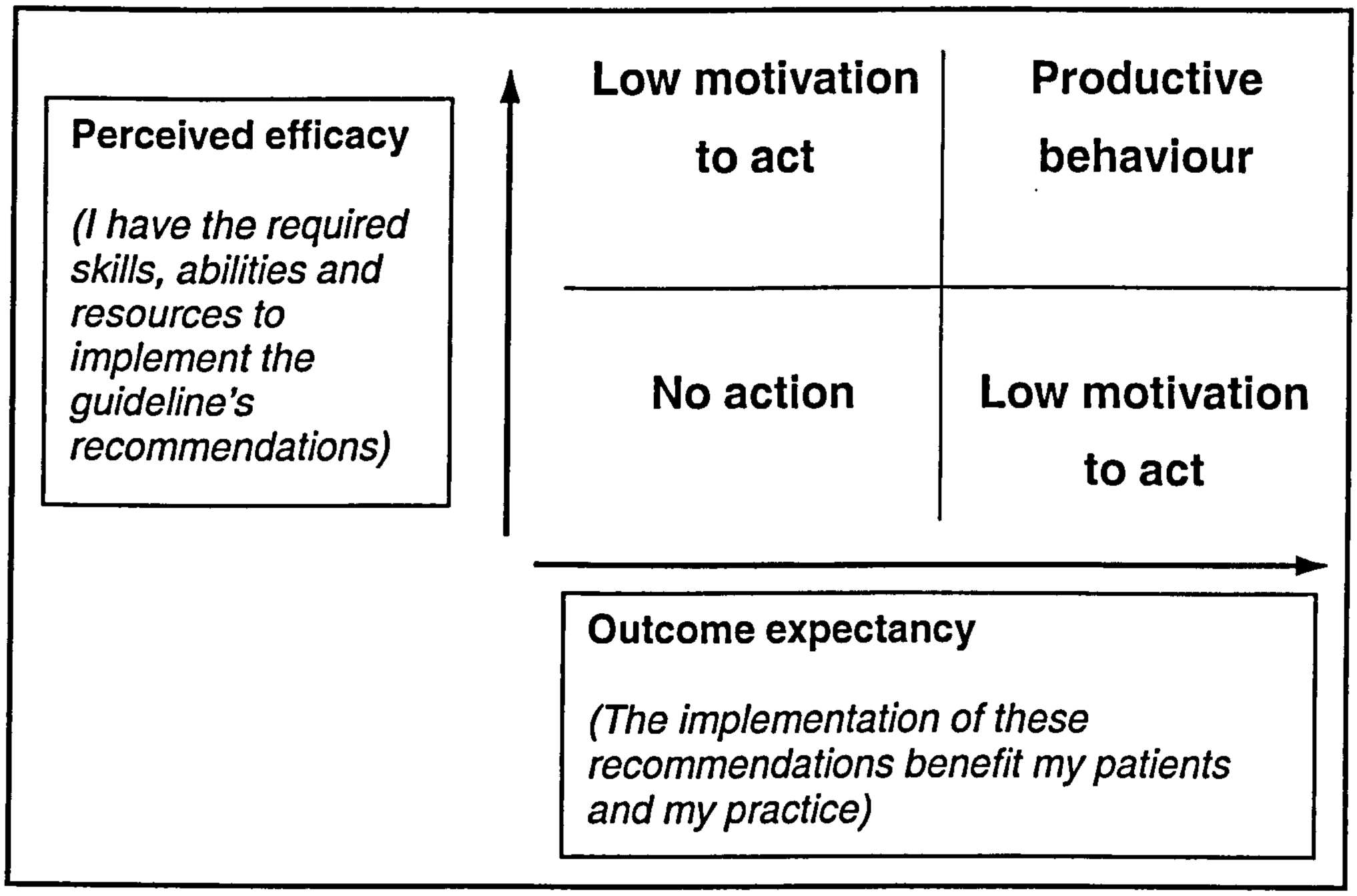
Stage-based theories maintain that behaviour changes happen as individual progresses through psychological stages of readiness to change (Norman and Conner, 1996). There are three main theories in this category: the transtheoretical model, the precaution adoption process model and the health action process approach (Riemsma et al, 2002). All three models maintain that for each behaviour people can be categorised into three groups: those who have not decided to change, those who have decided to change and those who are already changed (Riemsma et al, 2002). The transtheoretical model is by far the most frequently used model. Prochaska and DiClemente (1983) developed this individual level model of behaviour change in early 80s. It has been very popular among health promotion staff in the UK in the last decade (Riemsma et al, 2002). The 'stages of change', the 'processes of change' and the 'levels of change' are the three organising constructs of the model (DiClemente and Prochaska, 1998). In this model a continuum of behavioural stages is defined in which people progress towards the goal of behaviour change. In its earlier versions the model was a linear continuum of four stages. Further studies

resulted in the development of the model so that it comprised five elements or stages: pre-contemplation, contemplation, preparation, action and maintenance (Prochaska et al, 1992). In its current form the model does not imply that the progress through this continuum is linear. Individuals may go forward and backward during the process of behaviour change before they get to the maintenance stage. This process is called the spiral model of the stages of change (Prochaska et al, 1992). Prochaska et al (1992) integrated ten processes into the stages of change. Processes are resembled to the engines that facilitate movement between stages and are derived from several diverse theories (DiClemente and Prochaska, 1998). Levels of change are defined to take into account that the individuals with one problem behaviour have other problems that complicate the picture and the individuals may be at different stages of change for each problem (DiClemente and Prochaska, 1998). The model implies that interventions should be tailored to individuals' readiness for change (Riemsma et al, 2002).

Transtheoretical model has some limitations, mainly in its inability to explain what causes individuals to move between stages. It also has limited ability in explaining the effects of external factors (such as organisation and environment) on individual behaviour. Establishing the spiral movement of individuals between different stages requires long term longitudinal studies with repeated collection of data and complicated statistical analyses. Another limitation of this model, if applied to guideline implementation for prescribing, is that in most clinical scenarios it is likely to find GPs scattered at the later stages (namely preparation and action). At the theoretical level, the model has been criticised for artificial segmentation of natural continuum of change in attitudes and behaviours (Davidson, 1998). Davidson puts forward that pre-contemplation, contemplation and preparation are different levels of intention, and action and maintenance measure different levels of behaviour. He and others have argued the cut off points in distinguishing the stages are arbitrary (Davidson, 1998; Ogden, 2001). Sutton suggests the stages' definitions are logically flawed (Sutton, 2000, p 209). On the other hand, the model advocates hail the model's empirical success in a variety of health behaviours and different settings (Prochaska and DiClemente, 1998). A systematic review of randomized trials of stage-based interventions, however, concluded there was little evidence

to suggest these interventions were more successful that non-stage-based interventions or indeed the usual care (Riemsma et al, 2002). The transtheoretical model can help improving the effectiveness of interventions such as educational outreach and audit and feedback (Tables 2.1 and 3.3). Other stage theories might also help understanding of provider behaviour (Schmidt et al, 1990; Pathman et al, 1996).

FIGURE 3.1. SOCIAL COGNITIVE THEORY: INTERACTION BETWEEN OUTCOME EXPECTANCY AND PERCEIVED SELF-EFFICACY IN PRODUCING THE DESIRED BEHAVIOUR



Adapted from Community Sport and Recreation Initiatives Project, 2000.

Another psychological theory, social cognitive theory, is based on the principle of *reciprocal determinism* (Bandura, 1986; Baranowski et al, 2002). This principle suggests that three components of the individuals, their behaviours and their surrounding environment are constantly influencing each other (Baranowski et al, 2002). Social cognitive theory is content free (i.e. applicable to a wide range of behaviours; see 3.1.3) and has been defined as '*a theory of skill and competency management and cognitive behavioural control*' (Lewis, 2002, p266). The theory stems in efforts to explain human learning. A previous theory, operant learning, had argued that reward (or incentive) should be directly offered for learning to happen (Hicks, 2002, pp364-366). Albert

Bandura and others offered the alternative theory of social learning, which proposed learning occurs in a social context and is not conditional to the direct application of reward (Baranowski et al, 2002). The social cognitive theory was developed through Bandura's attempts to improve the theory of social learning (Bandura, 1977b; Baranowski et al, 2002). Social cognitive theory is a complex theory and includes several concepts (Bandura, 1998; Ogden, 2001; Baranowski et al, 2002). Self-efficacy is the prominent concept and the theory suggests that it is the main predictor of behaviour (Bandura, 1977a; Bandura, 1986). Introduction of self-efficacy was a major theoretical improvement in the understanding of human behaviour (Ajzen, 1998). Self-efficacy was later adopted in other major theories. Its addition to the theory of reasoned action (in form of perceived behavioural control) resulted in formation of the TPB.

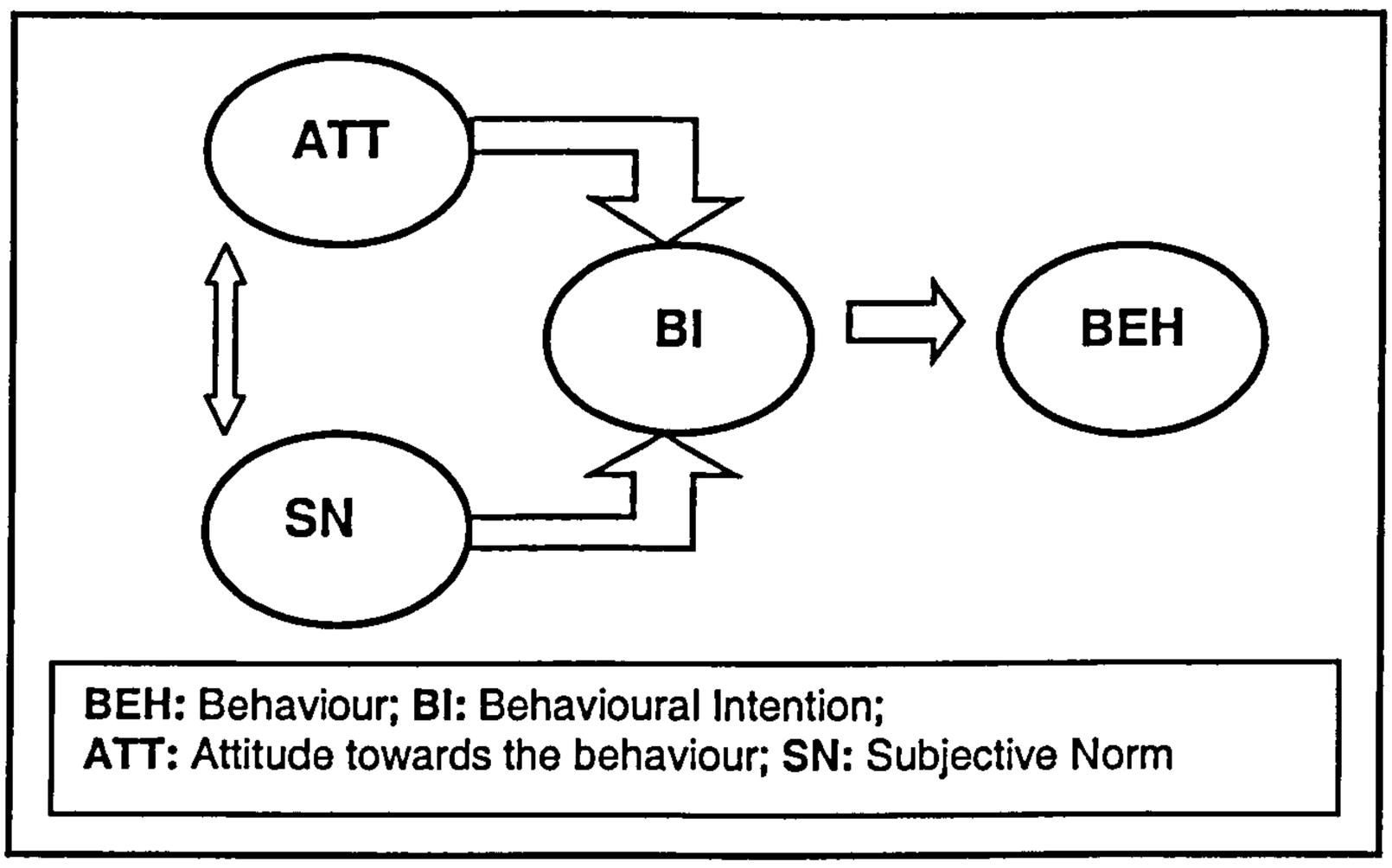
Social cognitive theory proposes that individuals' positive (negative) expectations of the outcome provide the incentive for (not) performing the behaviour. The individuals value the expected outcomes, and these values form the incentives to perform the behaviours (Baranowski et al, 2002). Both values and expectations filter through self-efficacy. Therefore if the perceived outcomes of engaging in behaviour are positive and the individual has a high level of self-efficacy, then the behaviour is likely to happen (Fig 3.1). Self-efficacy can be improved through social interactions, role models, training and simplifying the behaviour by reducing it to distinguishable parts. The concepts derived from the social cognitive theory and its predecessor, social learning theory, can help the understating of clinical behaviour change (Oxman et al, 1995; Davis et al, 1995). The concepts of the theory contribute to the understanding of several interventions, with varying levels of evidence of effectiveness (Table 2.1 and 3.3).

3.2.6. Theories of planned behaviour (TPB) and reasoned action

Theory of planned behaviour and its predecessor, theory of reasoned action, are among social cognition theories. Both theories are intended to explain volitional behaviours (Ajzen and Fishbein, 1980; Ajzen, 1985; Ajzen, 1988; Ajzen, 1991) and '*they appear to imply that individuals make behavioural*

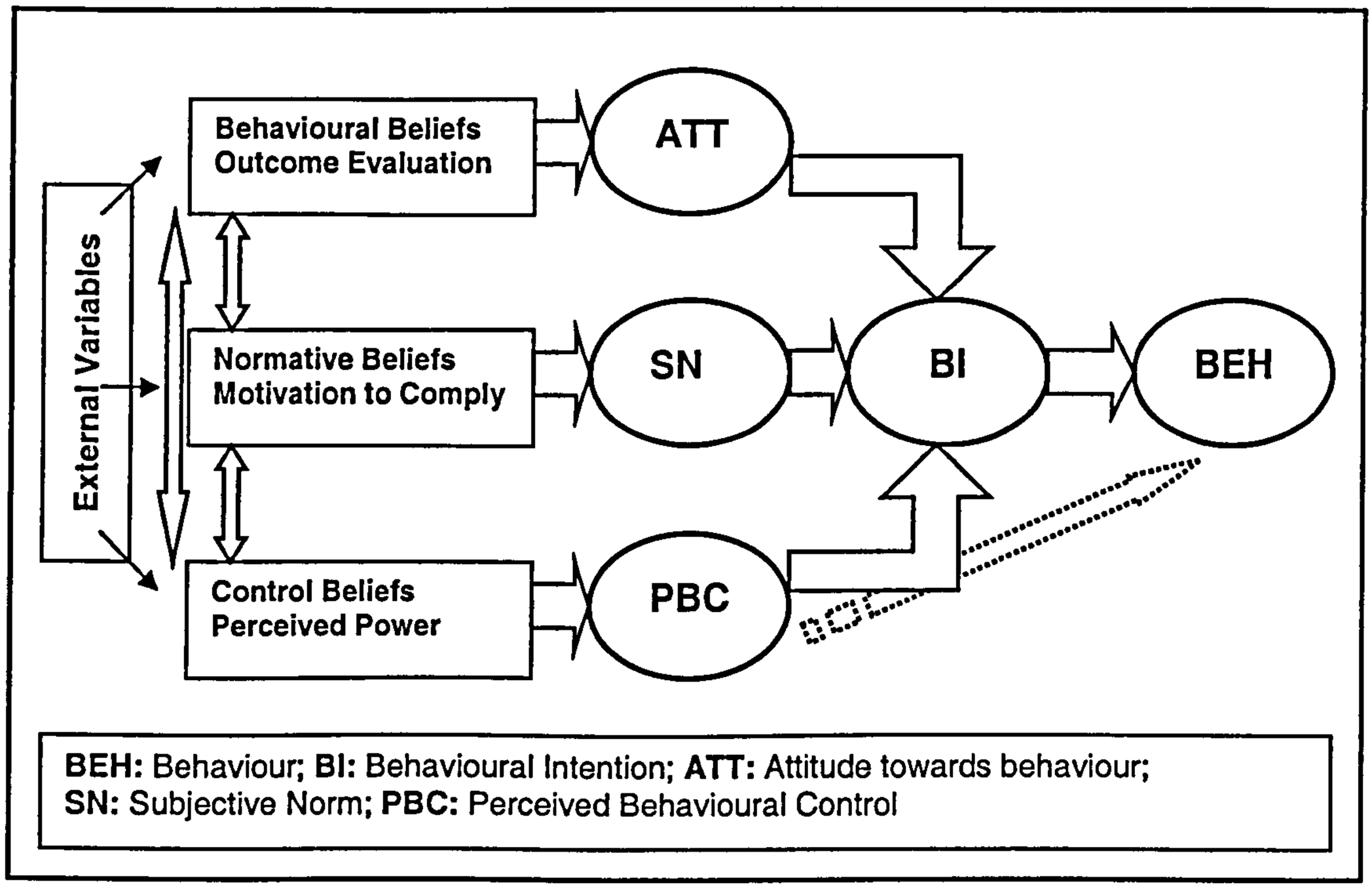
decisions based upon a careful consideration of available information' (Conner and Sparks, 1996, p 121). The above statement can be interpreted in a way that these theories especially the reasoned action are models of rational behaviour. Montano and Kasprzyk (2002) argue against this interpretation. They explain both theories assume individuals as *rational actors* so that individuals carefully process information before making behavioural decisions (Conner and Sparks, 1996). The information process may then change the underlying beliefs and through them the behaviour. That is regardless of whether the beliefs are rational or not. The assumption that individuals are rational actors is present in other social cognition theories (Norman and Conner, 1996) and in classical economic theories.

FIGURE 3.2. THEORY OF REASONED ACTION; SIMPLIFIED SUMMARY



Theory of reasoned action was developed by Ajzen and Fishbein (1980). The theory maintains that behavioural intentions are influenced by 'attitude towards behaviour' (attitude) and 'subjective norm' (or social cognition) (Fig 3.2). Attitude refers to the individual's overall evaluation of the behaviour. The emphasis on attitude towards *behaviour* is to distinguish it from attitude towards object and to prevent confusing different attitudes with each other (Montano and Kasprzyk, 2002). For example a GP's attitude towards asthma (object) may be different with his or her attitude towards prescribing a certain drug to treat asthma (behaviour). Theory of reasoned action maintains attitude and subjective norm are the determinants of behavioural intention (Conner and Sparks, 1996; Montano and Kasprzyk, 2002).

FIGURE 3.3. THEORY OF PLANNED BEHAVIOUR (TPB)



Theory of planned behaviour suggests that intention is determined by attitude towards the behaviour (attitude), perceived social or peer pressure (subjective norm) and perceived behavioural control (perceived control) (Fig 3.3). Perceived control is the person's perception of ability to perform the behaviour and control over carrying out the behaviour. It is very close to Bandura's (1977) concept of self-efficacy (Norman and Conner, 1996; see 3.2.5). It is also similar to Triandis's concept of *facilitating conditions*, which is about those personal or environmental characteristics that modify a person's behaviour regardless of person's behavioural intention (Montano and Kasprzyk, 2002). Hence, to put in jargon free context, TPB assumes that the individual makes his or her behavioural intentions based on three sets of beliefs and views: how the individual thinks of the outcome of the behaviour, how the individual thinks of the views of others about the behaviour (do important others expect or approve the behaviour) and whether the individual thinks to be capable of performing the behaviour (Marteau et al, 2002).

Intention is the cornerstone of the TPB as both predicted variable and predictor of behaviour (Fig 3.3). It should be directly measured using a set of

question items starting with 'I intend', 'I plan', 'I want' or 'I will' (Conner and Sparks, 1996; Ajzen, 2002b). The intention measured in this way is called the goal intention (Gollwitzer, 1999). According to the TPB, the strength of intention is the important predictor of behaviour. If there is a weak intention, the TPB suggests that altering attitude, subjective norm or perceived control may enhance intention. There is another psychological theory called 'implementation intentions' that suggests specific interventions for the strengthening of intention (Gollwitzer, 1999; Ajzen, 2002a). The theory asserts that behavioural intention is more likely to lead to behaviour if the intention is specified in terms of when, where and how (Gollwitzer, 1999; Gollwitzer and Oettingen, 2000). Therefore, Gollwitzer argues, by forming implementation intention the individuals commit themselves to certain reactions in given circumstances; and that is more than desiring a certain outcome in goal intentions. He adds that formation of implementation intentions helps the individual to '*switch from effortful control of their goal-directed behaviour ... to being automatically controlled by situational cues*' (Gollwitzer, 1999, p 495).

The theory proposes that attitude, subjective norm and perceived control are based on salient beliefs. Attitude arises from a set of beliefs about the behavioural consequences (behavioural beliefs) and evaluations of these consequences (outcome evaluations). Subjective norm is based on the individual's perceived views of others about the behaviour (normative beliefs) and the level of the individual's desire to adhere to the views of these people (motivation to comply). Likewise, perceived control is determined by individual's control beliefs and perceived power of those beliefs (Fig 3.3). According to the theory, internal and external factors might inhibit or facilitate formation of behaviour through control beliefs (Conner and Sparks, 1996). Internal factors for example are information, skills, emotions, abilities, and personal capabilities. Examples of external factors are opportunities, resources, social and organisational barriers and facilitators.

In a clear attempt to simplify the utilisation of the theory, the developers suggested that attitude towards behaviour was equal to the sum of products of 'behavioural beliefs' and 'outcome evaluations' (Conner and Sparks, 1996; Montano and Kasprzyk, 2002; Fig 3.4). For example for a GP to have a positive

attitude towards prescribing a specific drug for a given patient, she or he should positively value the expected change in patient's health and also believe that the drug is capable of achieving that change. Because of this theoretical reasoning TPB and other theories that utilise the same principle are referred to as expectancy-value theories (French and Hankins, 2003).

FIGURE 3.4. STATISTICAL PRESENTATION OF THE TPB

Regression models:

$BI = \beta_1 ATT + \beta_2 SN + \beta_3 PBC$

$BEH = \beta_1 BI + \beta_2 PBC$

Sums of products:

$ATT = \sum$ (behavioural belief × outcome evaluation)

$SN = \sum$ (normative belief × motivation to comply)

$PBC = \sum$ (control belief × perceived power)

BEH: Behaviour; **BI:** Behavioural intention; **ATT:** Attitude towards the behaviour; **SN:** Subjective norm; **PBC:** Perceived behavioural control

Similarly, subjective norm is the sum of products of normative beliefs and motivation to comply with those beliefs; and perceived control is the sum of products of control beliefs and their perceived powers (Conner and Sparks, 1996; Montano and Kasprzyk, 2002; Fig 3.4). For example, the feeling of social pressure is the result of the GP believing that her important others (e.g. peers) expect her to prescribe in given condition and the GP respects their views. This social pressure is again the overall pressure felt from all important others, which probably in this case are colleagues (clinical and non-clinical) and patients. A negative subjective norm can be observed because of the expectation of an untrustworthy or non-respected individual. Theory of planned behaviour proposes linear relationships between the model elements. Attitude, subjective norm and perceived control all have linear explanatory relationships with behavioural intention. In turn, intention and perceived control have linear relationship with behaviour. These relationships are best explained by two regression models (Fig 3.4).

The TPB maintains that attitude and behaviour need to be compatible, if attitude is to predict behaviour. This has been called the 'principle of compatibility' (Ajzen, 1988; Conner and Sparks, 1996). According to this principle, attitude and behaviour have four elements of action, target, context and time (Conner and Sparks, 1996; Montano and Kasprzyk, 2002). If attitude and behaviour are at the same level with respect to the elements, then the correspondence between attitude and behaviour will be the highest (Conner and Sparks, 1996). An example of this principle could be presented as a GP who prescribes anti-hypertensive medicines according to an evidence-based guideline (action), to reduce blood pressure of the patient (target), in his practice (context), tomorrow (time). This is only one presentation of a single behaviour of the GP. The behaviour could also be divided into separate behaviours (e.g. diagnosis, measurement, deciding on drug class, dose of drug, dealing with co-morbidities etc). Conner and Sparks (1996, p 134) state that the minimum specification of behaviour should include action and time frame. For health care providing behaviour of clinicians, specification of time as requested by the principle of compatibility may be a challenge. It is very difficult for a clinician to predict when they meet a patient that requires an activity that they intend to perform, especially in general practice. In theory the principle of compatibility is applicable to single behaviours (e.g. teeth brushing, prescribing a certain drug for certain condition) as well as general classes of behaviour (e.g. oral hygiene, guideline implementation; Conner and Sparks, 1996).

The main difference between TPB and theory of reasoned action is in inclusion of perceived control into the TPB. The other major difference is the way the relationship between behavioural intention and behaviour is devised. Reasoned action maintains that intention is the sole predictor of behaviour. However, TPB acknowledges that in many circumstances individuals perform (or refrain to perform) behaviours despite their intention. This has been presented through suggestion of a direct link between perceived control and the behaviour (dotted arrow in Fig 3.3). Therefore perceived control has a direct effect on behaviour without the mediating effect of behavioural intentions. In other words, perceived control moderates the effects of intention on behaviour (Baron and Kenny, 1986; Montano and Kasprzyk, 2002). The inclusion of perceived control into the model has extended the applicability of the theory

beyond easily performed volitional behaviours to those which require skills, resources or opportunities, and which are not considered to be within the domain of the applicability of theory of reasoned action (Conner and Sparks, 1996). A further advantage of adding perceived control is that theoretically the TPB will be able to take into account the individual's past experience. Previous experience of behaviour affects the individual's perceived behavioural control, and hence contributes to the formation of intention and behaviour. Thus, the TPB can be seen as the more appropriate of the two models with which to study provider behaviour since providers' clinical activities are influenced by a wide variety of external factors.

TABLE 3.2. EXAMPLES OF PUBLISHED STUDIES OF HEALTH CARE PROVIDERS USING TPB OR THEORY OF REASONED ACTION

Citation details	Method	Topic	Sample
Taylor et al, 1994	Expanded TRA; Survey	Screening mammography	85 general internists
Millstein, 1996	Longitudinal surveys	Offering STD prevention advice	765 doctors
Lambert et al, 1997	TRA; survey	Antibiotics prescribing	27 family doctors
Levin, 1999	Survey	Glove use	527 nurses and lab workers
Walker et al, 2001	Survey	Antibiotic prescribing	126 GPs
Watson and Myers, 2001	Survey	Glove use	103 nurses
McCarty et al, 2001	Survey	Smoking cessation	397 nurses in for hospitals
Meyer, 2002	Survey	Asking for assignments	92 nurse students

In order to utilise the TPB, investigators need to assess the views of a sample of respondents using qualitative approaches with the aim of identifying salient (or modal) behavioural, normative and control beliefs. Then the salient beliefs will be used for development of a theory based questionnaire which is then applied to the population under investigation (Conner and Sparks, 1996; Montano and Kasprzyk, 2002). Using salient beliefs is based on laboratory research on the structure and function of human memory in which only readily available elements of memory are used in decision making. The emphasis on salient beliefs also stems from earlier works on the relationship between attitudes and behaviour, where it is known as the 'specificity hypothesis'.

According to this hypothesis, specific attitudes towards the behaviour are more likely to predict the behaviour than general attitudes towards it (Raven and Rubin, 1983, pp 150-155). The individual is likely to have several beliefs towards a given behaviour. Normally only salient beliefs are processed in order to make behavioural intention. Using salient beliefs has clear practical advantages and is the common approach in TPB studies. Some researchers, however, have preferred to use subject-generated beliefs (Agnew, 2000). The need for preliminary qualitative interviews and feasibility studies are among criticisms of TPB (Rimer, 2002). These can increase the costs of performing such research and limit its applications. The other limitation stems from its strict restriction to three elements of attitude, subjective norm and perceived control. This may result in exclusion of other important psychological variables and limit the application of the theory (Rimer, 2002).

Other cognitive models based on expectancy-value structure have been previously applied to prescribing behaviour (Raisch, 1990a; Raisch, 1990b). 'Drug-choice model' was specifically developed to explain prescribing behaviour and maintained that practitioners' prescribing was determined by their belief towards the outcome expected from the drug and the value they assigned to it (Segal and Helper, 1982). Both TPB and reasoned action have been successfully applied to a variety of behaviours including smoking, alcohol consumption, food choice, exercise, sexual behaviours, recycling and dieting (Ajzen, 1991; Armitage and Conner, 2001; Hardeman et al, 2002). Until fairly recently, however, the majority of health related TPB research has been conducted to investigate the behaviour of patients or healthy populations, rather than health professionals. To date, a small number of studies have examined the utility of the TPB in investigating the behaviour of different categories of health professionals in a variety of settings (Table 3.2). The theory contributes to the understanding of several interventions through its concepts of subjective norm and perceived control (Table 2.1 and 3.3).

3.2.7. Other relevant theories

Several other theories are not discussed. One relevant theory is health education theory (Green et al, 1980; Moulding et al, 1999). Many aspects of this theory are covered in others that this chapter has introduced (Glanz et al, 2002a). The theory's main message is in its emphasis on the role of active involvement of the target population in changing behaviour. The theory supports participatory and interactive involvement of physicians in development, adaptation and adoption of clinical guidelines (Moulding et al, 1999). Learning theories and adult learning theories (Skinner, 1953; Cantillon and Jones, 1999; Kaufman, 2003) are also not discussed in detail. Different elements of these theories are incorporated in the theories introduced in the chapter, mainly social cognitive theory (3.2.5). Anyhow the classical theories of learning still provide stimulating messages for behaviour change. Laboratory research suggests that first-learned things generalise better over time and place (Bouton, 2000), therefore it is easier to introduce a new behaviour by clinical guidelines than changing a previously performed behaviour. The theory also suggests that lapses towards the old behaviour should be considered as a possibility and to be understood (Bouton, 2000). The other groups of theories which are not discussed in the chapter are those known as social marketing theories (Kotler and Roberto, 1989). The importance of following these techniques for guidelines implementation has been highlighted in the literature (Grol, 1997). Again many aspects of these approaches are covered in the theories already discussed (e.g. the diffusion of innovation). Health belief model (Becker, 1974; Rosenstock et al, 1994) is another theory that has been applied to provider behaviour (i.e. cancer screening; see Clasen et al, 1994). It is among the most popular health psychology theories (Glanz and Maddock, 2000; Glanz et al, 2002a). However, it shares many characteristics with other social cognition theories (e.g. TPB; Davidson, 1998; Quine et al, 2000). More importantly it is content specific (see 3.1.3) and application of content specific theories to provider behaviour may prove difficult.

3.3. Potentials of TPB for the study of GP prescribing

The selective review of the theories of behaviour change and their applications to health care providers demonstrated the intellectual depth that theoretical approaches could provide for health services research. Different academic disciplines provided opportunities to devise methods for tackling the great deal of variation still seen in prescribing and understanding the different reactions observed towards clinical guidelines. Using information summarised in this chapter, Table 2.1 was expanded to include relevant theories (Table 3.3). This attempt, however, did not overcome the main limitation, which was that most theories were not tested in empirical research. So the critical mass required for the careful application of the theories to provider behaviour was not attained. It should be noted that not all theories relevant to each intervention were mentioned in the table, for two reasons. First, the theories overlapped considerably, for example self-efficacy was part of different psychological theories. Hence if the contribution of the theories in understanding of the intervention was via this component, then mentioning one theory in the table considered sufficient. Second, the table was not meant to be comprehensive, but to link theories to interventions. Further detail was included in the text of the chapter. Table 3.3 included all the interventions that were identified and reported as the result of the overview of systematic review, except two interventions for which there were evidence of no effect (i.e. didactic CME and substitution model of inter-professional shared care). These were excluded.

Theories also indirectly contribute to health services research and quality improvement and that is by selective use of theories' components in the investigations and interventions. Many of these contributions are not fully acknowledged and are used only as practically known methods of changing behaviour. Despite this, it can be argued that assessing the theories of behaviour change in their totality is essential. If the theories are not assessed as a whole it is always difficult to relate success or failure of interventions or explanatory investigations to the theories.

TABLE 3.3. STRATEGIES FOR IMPROVING PRIMARY CARE PRESCRIBING AND THEORIES THAT MAY EXPLAIN VARIATION IN THEIR UPTAKE

		Intervention	Relevant theories	Effect	Cost	Durability
Voluntary/ Internal motivation	**Competence oriented**	CME (interactive)	Adult learning; TPB	++	Medium	?
		IP education	Adult learning	?	Medium	?
		Mailed printed material	Social influence; TPB	+/0	Low	?Short
		Mailed national warnings [$]	Social influence	+	Low	Medium to long
		Participatory guideline devel.	Social influence; TPB; health education	+	High	?Medium to long
		IP shared care (consult.-liaison)	Health education; TPB	+/0	?High	?
		Mass media	Diffusion of innovation; TPB; social influence	+	Low	?Medium
	Performance oriented	Audit and feedback[$]	TPB; social influence; stages of change	+	Low to medium	?
		Reminder systems	TPB	?+	Low to medium	Short
		Educational outreach	Diffusion of innovation; social influence; stages of change; TPB	?++	Low to high	Short to long
Voluntary/ External motivation	**Social influence**	Peer review	Social influence; TPB; Diffusion of innovation	?	?Medium to high	?
		Patient mediated	Social influence; TPB	+	Low to medium	?
		Local opinion leaders	Diffusion of innovation; Leadership; TPB	?	?Medium to high	?
		CQI [$$]	Management theories	?	?Medium to high	?
	Physical support	Practice support	Management theories; TPB	?	?	?
		Essential drugs programmes [%]	Management theories	?+	?	?
	Financial incentives	Financial incentives	Economic theories/ social influence	?+/-	Low to high	?Short to medium
Non-voluntary		Reimbursement and budgetary policies	Economic theories; TPB; social influence	?+/-	?	?
		Rules, obligations	Economic and management theories; social ecology	?	?	?
		Restricted formulary	Economic theories; social ecology	+	?	?Medium to long
		De-registration / reclassification	Economic theories; social ecology	?+/-	?	?

* Adapted from Table 2.1, Grol (1992) and Wensing and Grol (1994). IP: Inter-professional, ++: strong evidence suggests positive (intended) effects, +: limited evidence suggests positive (intended) effects, +/0: variable effectiveness, ?: no evidence of effect, ?+: no evidence of effect, however less reliable evidence suggests positive effects, ?+/-: no evidence of effect/ likelihood of positive and negative (intended and unintended) effects, $ it usually also has an element of social influence, % it usually incorporates competence oriented approaches, $$ it usually has elements of physical support and competence oriented approaches.

Theory of planned behaviour provides interesting avenues to explain variation in provider behaviour. First, TPB, unlike most other theories presented in this chapter, satisfies Kerlinger's strict definition of theory (Table 3.1; see also Rimer, 2002). That is to say, the theoretical components of the theory are properly codified and the theory suggests methods of assessing the relationship between the components. Therefore, it is possible to test TPB in empirical research. Second, the theory is a thoughtful elaboration of the behaviours of rational actors. Hence, its components contribute to the understanding of variation in effectiveness of several interventions to improve prescribing as presented in Table 3.3. Third, it has been applied in a few circumstances to explain the variation in intention and occasionally the reported behaviour of clinicians (Table 3.2). Although the reported applications of the TPB to provider behaviour are limited, they demonstrate general support for the theory as a potentially useful tool for understanding provider behaviour. Fourth, none of previous TPB studies included objective measures of clinician behaviour and instead relied on reported behaviour. This limited the reliability and validity of the reported goodness-of-fit of the models. It is therefore essential to assess TPB using actual behavioural measures before making conclusions. Fifth, the theory has never been used before to explain variation in prescribing intentions and prescribing outcomes in relation to clinical guidelines. Sixth, there was a strong view among some leading health services researchers and health psychologists that TPB might provide part of the answer to the understanding of provider behaviour. This was particularly encouraging as it was assumed that the results of assessing TPB would be of immediate use to the scientific community as well as policy-makers. Using all these arguments, it was concluded to assess the validity of TPB as a theoretical tool for the understanding of GP prescribing in accordance to clinical guidelines. Nonetheless, there was no claim that the TPB was the only suitable theory in need of being assessed, nor that it was the only theory that might help provider behaviour change and quality improvement initiatives.

The following chapters of the thesis report the findings of the primary studies. Chapter 4 reports the findings of the qualitative study of GPs' views on implementation of guidelines to improve prescribing. The qualitative study was

conducted to achieve two goals: one to gain better understanding of variations in implementation of clinical guidelines, and second to identify GPs modal beliefs in order to test the TPB. The main results of the formal assessments of TPB using methods recommended by its advocates (i.e. questionnaire surveys) are reported in Chapter 7.

Doctors are men who prescribe medicines of which they know little, to cure diseases of which they know less, in human beings of whom they know nothing.

Voltaire (1694 - 1778)

Chapter 4. The complexities of changing primary care prescribing: a qualitative study

4.1. Introduction

This chapter presents the results of a qualitative study of GPs and primary care academics to understand the variations observed in the implementation of clinical guidelines' prescribing recommendations. It reports the results of the analysis conducted using the framework method. It also gives details of the thematic framework developed through the analysis and the insights that the study adds to the findings of the previous stage of the project.

4.1.1. Background

In the last two decades improving quality of prescribing attracted a lot of interests, not least because important fraction of health care resources were (and increasingly are) spent on it (Donaldson and Donaldson, 2003). Despite the availability of evidence, changes in prescribing are delayed and require planning and investments. For example, clear-cut messages for antibiotic prescribing are still to be implemented (McEwen et al, 2003) and although there are numerous studies on how to improve prescribing, but still all reasons behind 'nonscientific prescribing' are not known (Kumar et al, 2003). While there was evidence that neither demographic and social characteristics (Soumerai and

Avorn, 1987) nor organisational culture (Dowswell et al, 2001) predicted physician prescribing, few studies tested the theoretical models that acknowledged psychological and organisational processes which preceded behaviour (Marteau et al, 2002). Focus on GPs' attitudes and personal beliefs may provide a better insight of prescribing. The importance of identifying underlying personal reasons for nonscientific prescribing is documented in the literature. A multi-centre trial studied expressed reasons physicians' nonscientific prescribing' and concluded that '*greater attention must be paid to physicians' attitudes and motivations concerning suboptimal prescribing if programmes are to succeed in replacing these practices with more rational clinical decision making*' (Schwartz et al, 1989, p 577).

4.1.2. What role for qualitative studies?

Underlying beliefs are better captured through qualitative studies (Comaroff, 1976). Qualitative studies should play important roles in the assessment of health technologies (Black, 1994; Leys, 2003) particularly in general practice (Murphy and Mattson, 1992; Jaye, 2002). Interviews could be particularly useful in identification of barriers to guideline implementation and in understanding how those barriers affect individual physicians (Pathman et al, 1996). Surveys then would be very useful in quantifying these qualitative findings in larger scale (see Cabana et al, 2001 as an example in asthma care). A few qualitative studies have focused on prescribing. A 1996 study of change in prescribing found that GPs freely spoke of their prescribing and were ready to offer embarrassing observations during the interviews. The study demonstrated that interviews were useful tools of studying prescribing in primary care. The study identified three models of change in prescribing: accumulation model, challenge model and continuity model (Armstrong et al, 1996). Another British study of GPs' and consultants' reasons for change in clinical practice found that education was important in changing prescribing (Allery et al, 1997). A qualitative study of 24 Scottish GPs assessed statins prescribing using semi-structured interviews. It concluded that GPs rarely critically appraised trial results, but evaluated the results in terms of their social and economic

implications (Fairhurst and Huby, 1998). It suggested that compiled sources of evidence (e.g. evidence-based guidelines) could play a useful role. It implied that GPs were likely to rely on trustworthiness of the source of the guidelines, and that GPs considered the implications of implementing clinical guidelines. Another qualitative study of prescribing, performed on seventeen GPs, tried to identify variables that explained prescribing variance (Carthy et al, 2000) and concluded that consultant's prescribing was among those variables.

Salisbury and colleagues (1998) used qualitative and quantitative methods to audit three changes in prescribing (use of warfarin or aspirin for atrial fibrillation, angiotensin converting enzyme inhibitors for heart failure and substitution of trimethoprim for co-trimoxazole) and concluded there were no effects from clinical protocols. They reported positive views of participants towards clinical protocols, but found that many did not use them for these key changes. They compared practices that used protocols for heart failure treatment with those that did not, and found no statistical difference (Salisbury et al, 1998). Their conclusion of no effect may not be correct, owing to the potential lack of statistical power. Observing conflicting results in terms of guidelines implementation, Grol et al performed an observational study to identify those attributes of clinical guidelines that influenced implementation (Grol et al, 1998). The Dutch study concluded that '*specific attribute of clinical practice guidelines determine whether they are used in practice*' (ibid, p 861).

There had been further qualitative studies to elicit broad issues such as 'GPs views on use of guidelines' or 'their attitudes towards evidence-based medicine in general' (Langley et al, 1998; Mayer and Piterman, 1999). These issues are too broad to be covered in single studies, as the researchers would be more likely to get general statements in response. In contrast there had been disease (or clinical area) specific studies of guideline impacts on practice. For example, a qualitative study of patients suggested that patient willingness to take warfarin played a major role in the implementation of atrial fibrillation clinical guidelines (Howitt and Armstrong, 1999). Another recent example was qualitative study of how to implement hypertension guideline recommendations for the elderly patients (Cranney et al, 2001).

The overview of systematic reviews (Chapter 2) demonstrated that no single intervention or combinations of interventions were successful for clinical guidelines implementation across settings or contexts. Therefore it was decided to assess the views of GPs about the factors that facilitated or inhibited clinical guideline implementation. To ensure capture of important variables of changing prescribing in primary care, it was decided to focus on a small set of different clinical conditions. This selective approach could provide the opportunity of comparing different clinical conditions, clinical guidelines and dissemination strategies. The results of qualitative study were also considered important for informing the next phase of this PhD project, namely the development of the questionnaires to assess the ability of TPB in explaining variation in prescribing.

4.1.3. Objectives

The qualitative study was aimed to improve understanding of the findings of the previous phase of the project, to be used for the development of appropriate research tools for the next phase of the project and to identify key themes for the implementation of clinical guideline prescribing recommendations in primary care. Specifically the qualitative study objectives were:

- To explore key themes for the implementation of clinical guideline prescribing recommendations in primary care
- To explore GPs' attitudes towards clinical guidelines; and to identify salient attitudes and beliefs about prescribing in accordance with guidelines to develop TPB survey questionnaires for the next phase of the project
- To explore barriers to and facilitators of implementation, and to explain differences in implementation
- To suggest approaches and strategies to improve implementation of clinical guideline prescribing recommendations in primary care and to improve the understanding of variations observed in the previous phase of the project (i.e. the overview of systematic reviews)
- To identify two appropriate clinical conditions to be used in the next phase of the project (i.e. the TPB surveys of GPs)

4.2. Methods

4.2.1. Setting

The study was conducted in British primary care with emphasis on England. Respondents were from two former Primary Care Groups located in North East of England as well as individuals from other parts of England, Wales and Scotland.

4.2.2. Participants

A purposeful sample of 25 participants was interviewed. Twelve participants were GPs (one trainee, ten GPs from practices in York and Selby and one GP from Scotland). Three were female. The participants worked in partnerships of different sizes ranging from two partners (one fulltime and one part-time) to ten partners. Another sample of thirteen GPs and academics of primary care from other parts of Great Britain were invited to take part in the interviews. These interviewees were categorised as academics of primary care. In this group nine respondents were practising as GPs as well as their academic appointment, one was formerly a GP and three were non-GP academics of primary care. In this group two interviewees were from Scotland, two from Wales, one from Manchester, one from London and seven from North and East of England. Two were female. Amongst the academic group seven held chairs.

GPs were invited who had interest in or experience of clinical guideline implementation, whether or not they agreed with the guidelines. They were identified in consultation with a local GP trainer, a University of York health services researcher and the interviewees. Academics of primary care were invited to ensure a more widespread account of the issue. The criteria for choosing this group were extended experience in development, implementation

or evaluation of clinical guidelines. The participants were identified in consultation with two University of York health services researchers.

Eighteen GPs were invited of whom twelve were interviewed. Sixteen primary care academics were invited of whom thirteen were interviewed. One GP declined the interview because his practice was flooded and they were involved in sorting out the problem. My vehicle broke down on my way to another arranged interview. Attempts for re-arranging it were unsuccessful. Another GP agreed to be interviewed and then declined owing to other commitments. Others provided no reason for declining. The overall response rate was 74%.

4.2.3. Clinical conditions

This qualitative study deliberately focused on five clinical conditions for which there were evidence-based clinical guidelines and recommendations and prescribing was important part of managing the disease. The clinical conditions were asthma (in adults and children older than five years), use of statins for prevention of coronary heart disease (CHD), epilepsy, menorrhagia and depression. The conditions provided a suitable combination of clinical guidelines: high awareness with complicated prescribing (asthma), high awareness with less complicated prescribing decisions (statins for CHD prevention), medium awareness with focus on diagnosis rather than treatment (depression) and conditions for which guidelines were not well publicised in primary care and GPs had varying degrees of responsibilities (menorrhagia and epilepsy). All those clinical conditions resembled each other in the fact that GPs prescribed medicines for their treatment at some stage of the disease. Also for all of them there were recognised clinical guidelines with prescribing recommendations.

4.2.4. Interviews

The interviews were conducted in late 2000 and early 2001. Twenty-five GPs and academics of primary care took part in the interviews, of which, eleven were face-to-face (with those living or practising in or around York) and fourteen were telephone interviews. Twenty-five in-depth or semi-structured interviews (Britten, 1995; Morse and Field, 1996) were conducted lasting 25 to 45 minutes. The interviews were planned so that there were minimal chances of disruption (Britten, 1995). The three starting interviews were face-to-face and in-depth and were conducted with one GP, one GP trainee and one academic with interest in guideline implementation and prescribing in primary care. The purpose of these exploratory interviews was to provide a better understanding of the context and also to provide a suitable set of questions for the semi-structured interviews. They were also used to identify five clinical conditions with different characteristics to focus on during the semi-structured interviews.

The author conducted all the interviews following the recommendations of Morse and Field (1996) and Ajzen (2002). Therefore the interview questions were devised so that they captured salient beliefs of the GPs, which were to be used for the development of the TPB survey questionnaires (i.e. next phase of the project). The study objectives were stated at the start of each interview. The interviewees were consented to tape-record the interviews for the purpose of this study with the ethical intention of respecting the confidentiality of the interviewees. In semi-structured interviews, about ten open-ended questions were asked (Appendix IV). The first two questions were general and asked about the respondents' experience of useful and non-useful clinical guidelines. The following questions focused on the five clinical conditions (asthma, depression, epilepsy, menorrhagia and statins for CHD prevention). The next set of questions asked about influential people in clinical guideline implementation, and other barriers to and facilitators of implementation. In the end, the interviewees were invited to add any extra points or comments. All of the interviews were transcribed for the analysis.

4.2.5. Thematic framework and qualitative analysis

For analysis, we followed the 'framework' method (Ritchie and Spencer, 1994; Ritchie et al, 2003). This method had two advantages. First, no specialised computer software package was required. Here the Microsoft Word software package was used, although the job could be done using spreadsheet packages as well. Second, the method had been specifically developed for the analysis of qualitative data for policy-oriented projects. Health services researchers had used the method successfully (e.g. Griffiths et al, 2001). Ritchie and Spencer (1994) explained that '*qualitative data analysis is essentially about detection, and the tasks of defining, categorising, theorising, explaining, exploring and mapping are fundamental to the analyst's role*' (p 176). The framework was found a very useful method to fulfil these tasks.

The framework has five broad steps. The method consists of 'familiarisation', 'identifying a thematic framework', 'indexing', 'charting' and 'mapping and interpretation'. The steps were followed very closely. Ritchie and Spencer (1994) recommended that the analysis should start with data familiarisation, especially where not all interviews were performed by the person responsible for the analysis. This was not a concern in this study, because all the interviews as well as transcribing were conducted by one person. Nonetheless all the interview tapes were listened to again, and a contact and content summary form was developed and filled out for each interview during the process (Miles and Huberman, 1994, p 54). The initial thematic framework was developed based on the interviews, prior thoughts and literature review (Chapters 2 and 3). Theoretical knowledge, mainly the TPB (Ajzen, 1991), was exploited for the development of the thematic framework (deductive approach; see Pope et al, 2000). Deliberate attempts were made to include identifiable variables that corresponded to the main elements of TPB (i.e. attitude towards the behaviour, subjective norm and perceived behavioural control) in the framework. A preliminary framework was developed and then discussed in a series of iterative meetings between the researcher and his advisor. Then the thematic framework was checked against the interviews through repeating the familiarisation process. The themes were updated in the process of the analysis (Pope et al, 2000). The initial framework contained nine themes which were

reduced to seven as the analysis developed. The process of refining the themes followed a non-strict adoption of grounded theory (Barbour, 2001). Where theoretical reasoning and literature were not supported by data, superiority was given to the data (inductive approach; see Pope et al, 2000).

The transcribed text was 'indexed' using the codes relating to the themes and sub-themes of the thematic framework. Sections of data were indexed with one or more codes (cross indexing) wherever appropriate (Pope et al, 2000). For each 'theme' one table was produced in Microsoft Word. Each table row was assigned to one interviewee identified with appropriate code. Table columns were assigned to the sub-themes. The tables were then attached together to produce the analysis 'chart'. The final chart was a large table of many cells, in which each cell corresponded to the views expressed by one interview coded for one sub-theme. The data extracts were then 'cut and pasted' from the indexed transcribed texts to appropriate cells on the chart. The chart was then printed on A3-sized papers for the analysis. Therefore it was possible to compare the views expressed by one interviewee across different themes and sub-themes (by looking across the rows), and to compare the views of different interviewees for each theme (by looking across the columns). The main transcription files were consulted in the course of the analysis and further extracts were added to the chart whenever it was necessary. The chart was used to interpret the data by comparing the data extracts for individuals and sub-themes. The relationships between the sub-themes and themes were also investigated.

When quoting from the interviews, each interviewee was represented with a code. GPs' codes started with the letter 'P' and academics of general practice codes started with the letter 'A'. In this report, curly brackets at the end of italic statements refer to the interviewee codes.

4.2.6. Ethical considerations

The study was conducted with careful consideration of ethical issues. The participants agreed to take part in the study after receiving letters or emails that contained information about the purpose and methods of the study. The time, venue and method of the interviews (face-to-face or telephone) were mutually agreed upon. The purpose of the interviews and the study objectives were reiterated at the beginning of each interview. All the participants consented to their interviews to be tape recorded and used for the purpose this study. Tape recordings were appropriately coded and stored securely. Transcribed information were coded and stored securely. The participants were assured that the research reports would not contain any specific details which would allow the identification of specific individuals or practices. Local Research Ethics Committee approvals were not sought as at the time of data collection interviewing health professionals did not require ethics approval. No honorarium was offered to the participants.

4.3. Findings

4.3.1. Description of the issues discussed in the interviews

Clinical guideline influence on GP prescribing was the focus of the interviews. However, other behaviours such as diagnosis and (cancer) screening were also discussed, although the interviewees were not asked about those. The first question asked the interviewees to name any clinical guideline that they recalled as useful in improving prescribing. In response, there was a widespread recognition that the British Thoracic Society (BTS) asthma guidelines were useful (British Thoracic Society, 1993b; British Thoracic Society and National Asthma Campaign, 1997a; British Thoracic Society and National Asthma Campaign, 1997b). This was followed by guidelines for primary and secondary prevention of coronary heart disease. The interviewees referred to Sheffield (Hag et al, 1995; Ramsay et al, 1996) or New Zealand (Jackson, 2000) risk tables, or to the Joint British Recommendations (British Cardiac Society et al, 1998; British Cardiac Society et al, 2000). As mentioned in 4.2.3 asthma, primary and secondary prevention of coronary heart disease, depression, epilepsy and menorrhagia were directly covered by the interview questions. During the interviews other clinical conditions such as COPD, osteoporosis, UTI, influenza, drug addiction, gastroenteritis and diabetes were also discussed although there were no specific questions about those diseases.

4.3.2. Themes, sub-themes and items

As a result of the analysis seven themes and thirty sub-themes were identified. It was aimed to limit the number of themes and sub-themes to prevent confusion and increase the coherence of the analysis. The themes and sub-themes were important issues for the implementation of clinical guidelines to change GPs' prescribing. Seventy-seven items were categorised within the sub-themes (Table 4.1).

TABLE 4.1. THEMATIC FRAMEWORK FOR QUALITATIVE ANALYSIS OF THE INTERVIEWS: THEMES, SUB-THEMES AND ITEMS

Theme I: Credibility of content of clinical guideline	
I.1 Evidence	
Evidence-based clinical guideline	Non evidence-based clinical guidelines
Strong evidence is not available	
I.2. Change	
Reflecting GPs normal practice	Change in clinical guideline message
Clinical guideline for areas that are changing	
I.3. Flexibility	
Each patient is unique	Many patients have more than one problem
Clinical guideline allows flexibility	Clinical guideline is inflexible
Having negative effect on patient-doctor relationship	Having positive effect on patient-doctor relationship
Theme II. Credibility of source of clinical guideline	
II.1. Reputable bodies	
Clinical guideline is from a reputable body	
II.2. Secondary care as reputable body	
Guideline produced by secondary care	GPs' ideas haven't been represented in it
Secondary care do not understand GPs	It's far from general practice realities
II.3. National bodies as credible sources	
NICE and SIGN are influential	NICE guidelines are not helpful
NSFs	
II.4. Pharmaceutical companies	
Drug companies and their representatives	
II.5. Publishing guidelines	
Publishing in respected sources	Inclusion in the BNF
Theme III. Presentation of clinical guidelines	
III.1. Simplicity	
Complicated message	Simplicity of clinical guideline
III.2. Systematic presentation	
Stepped approach	
Theme IV. Influential people in implementation	
IV.1. Patients	
Patients as facilitators	Patients as barriers
Patient compliance (or preference)	
IV.2. Consultants	
Consultants are positively influential	Consultants are neutral
Consultants are negatively influential	
IV.3. GP colleagues	
GP colleagues within practice	GP colleagues outside practice
GP colleagues are neutral or negatively influential	
IV.4. Practice nurses and primary care team	
Practice nurses	Primary care team as a whole
Nurses follow clinical guidelines (implementation rely on nurses)	Professional boundaries between doctors and nurses (nurse led clinics)
IV.5. PCOs, pharmacists and prescribing advisers	
PCGs are influential	PCGs are neutral
Pharmacists, prescribing advisers	
IV.6. Drug companies and reps	
Drug companies and their representatives	

TABLE 4.1. THEMATIC FRAMEWORK FOR QUALITATIVE ANALYSIS OF THE INTERVIEWS: THEMES, SUB-THEMES AND ITEMS

Theme V: Organisational factors	
V.1. Practice characteristics	
Dispensing or training status	Having clinics
V.2. Information technology	
Computerised system (clinical guideline included)	Computerised practice
V.3. Availability of required resources	
Necessary hardware unavailable to GPs	
V.4. Time, workload and information overload	
Lack of time; increased workload	Information overload
Having a single national guideline for problem	
V.5. Cost and expenditure	
Clinical guideline justifies extra costs and prevents unreasonable costs	Increased costs and expenditure
Theme VI. Disease characteristics	
VI.1. Treatment is secondary care based	
Treatment is mostly secondary care based	
VI.2. Difficulty of diagnosis	
Diagnosis is the major problem not treatment	
VI.3. Rare or 'simple' disease	
Being familiar with clinical guideline message but not the guideline itself	Non-remembering the clinical guideline
Theme VII. Dissemination strategy	
VII.1. Planning Implementation	
Receiving clinical guideline through the post is not useful	Lack or weakness of implementation programme
Implementation programme	
VII.2. Can implementation be successful?	
Difficult to relate the achievements to guidelines	Clinical guideline is rubbish or useless or not helpful
Misunderstanding clinical guideline message	Clinical guideline changed my practice
VII.3. When do GPs welcome new guidelines? Perceived need, past experience, and knowledge	
Positive experience with clinical guideline	Having a perceived need to a clinical guideline
Clinical guideline increases the knowledge base	Perceived as non-useful in first contact
VII.4. Ownership – local versus national guideline	
Practice- based (or locally adapted) guideline	Being on the guideline development team
VII.5. Enforced implementation; medico-legal issues	
Medico-legal issues	Audit and clinical governance
Enforcing clinical guidelines by quality markers and indicators	Dissemination perceived insulting or imposing
VII.6. Supporting implementation and implementation cost	
Financial incentives and resources	No financial incentive
Repeating clinical guideline message	Cost-containing motives for dissemination

The first three themes referred to the important characteristics of clinical guidelines that influenced the implementation of prescribing recommendations. The themes were 'credibility of content', 'credibility of source' and 'presentation' (Table 4.1).

4.3.3. Theme I: Credibility of content of clinical guideline

Three characteristics of clinical guidelines were identified that together helped GPs to form their judgement about the credibility of the content of prescribing recommendations. Those were: evidence that guideline was based on; relationship of guideline to change in practice and change in guideline itself; and flexibility of the message and its applicability to a wide range of primary care patients.

I.1. Evidence

The interviewees strongly insisted on the importance of inclusion of evidence-based recommendations in clinical guidelines. In their view, being 'evidence-based' was a major characteristic of any acceptable guideline. "*You got to have trust and faith in the fact that [clinical guidelines] have been correctly researched*" (P2). One interviewee, referring to the current BTS asthma guidelines, said: "*they are actually based on evidence of effectiveness from things like the Cochrane database and randomised controlled trials. ... they've actually gone through the right rigorous process of guideline formulation. So that's why I'm happy just to use them as a blueprint*" (A11). He also added that weak evidence-base resulted in non-adherence to guidelines: "*where the evidence isn't excellent, it's obviously more difficult to use a guideline*" (A11). Another interviewee insisted that "*I'm not saying I use [CHD prevention guideline] as necessarily as they've laid out, because hypercholesterolaemia is an area that I think has been poorly researched and we are at the moment working on a baseline that has never been established*" (P10).

However, not all interviewees defined the best evidence as it was defined in classic evidence-based medicine (Sackett et al, 1996). Some types of evidence did not seem so reliable to GPs. In their view, 'evidence' should come from trials of primary care patients, including almost all primary care patients, with comprehensive characteristics, such as the presence of co-morbidity, older patients and different presentations of a disease. One interviewee explained

that guidelines had to be "*based on good evidence, ... a number of trials, some sort of meta-analysis. The other good evidence would be pragmatic trial based in primary care. But unfortunately a lot of them aren't. So a lot of people I'm dealing with would have been excluded from the explanatory trials. Which means that I haven't got any good evidence to support my decision making*" (A9). They admitted that this sort of evidence was not always available and added perhaps that was why the guidelines were not so useful to them. "*Nobody really devised it up so that you can see the mild and moderate forms of depression that we diagnose in primary care so often, [to say] which [antidepressants] are best for those different conditions*" (A11). "*... who's to say what percentage of patients should be on beta-blocker after MI? Because nobody trials our patients*" (A9). In this context, if the GPs perceived 'evidence' as irrelevant or unconvincing, they concluded there was no evidence available.

Apparently some GPs adopted a passive approach to acquisition of knowledge derived from emerging evidence. '*They*' (GPs) should '*become aware*' of the evidence behind recommendations. In this approach, the responsibility of distributing evidence and convincing GPs fell upon disseminators, researchers and policy-makers and not the practitioners. "*A couple of years ago GPs were apparently under scrutiny because they were referring too many people for PSA [Prostate Specific Antigen], where now the hospital expects to see, well ... yes, do PSAs on everybody. I'm not aware of treatment having changed and I'm not aware of the evidence of how to treat CA prostate having changed*" (P5). This statement could be read in two ways: it might mean that 'I am aware that there is no evidence to support the change in practice'. On the other hand, it might imply that 'because I am not aware of any change in evidence, therefore there has been no change'.

Regardless of all the emphasis on evidence as a major characteristic of an implementable guideline, there were important exceptions. Few referred to the BTS asthma guidelines as evidence-based and most considered the BTS guidelines as not necessarily based on evidence, at least in their early versions. However, the interviewees believed that the BTS asthma guidelines were largely adopted in general practice. They mentioned other (poorly defined) element as the underlying reason for this success, practicality "*the fact that*

some of [BTS] wasn't evidence-based didn't seem to make a lot of difference. It was a practical set of guidelines" (A13). It seemed that a practical guideline could apparently overcome its shortcomings in evidence. Although in the literature being 'evidence-based' was defined as the major characteristic of clinical guidelines, to practitioner evidence was only one element among others. On the other hand supposedly evidence-based asthma guidelines (North of England Asthma Guideline Development Group, 1996; Centre for Health Services Research, 1996; North of England Evidence Based Guideline Development Project, 2001) might have failed to appeal to GPs. *"I think if you send me a copy of North of England guidelines for GPs, I will be very surprised if I read them, I will be very surprised if I refer to them, because I think I'm pretty good in asthma care, but it could be possible then to start a sort of cycle or build a system around me ... and support me to use the guidelines ...*" (A10).

For a variety of reasons, GPs perceived many guidelines as unhelpful or useless. Most importantly it was due to the perceived lack of evidence behind a guideline. "*Am I thinking the guidelines are rubbish? And I do. Quite a lot of guidelines that come ... and I do quick reading, I say I don't accept that. ... if I don't see strong evidence linked*" (P5). Failure of clinical guidelines was not due to GPs but the guidelines themselves. "*There is a common view that the failure of guidelines is simply due to a resistance or the stupidity on the part of general practitioners, which I think it is very unreasonable. It's mostly because the guideline is simply not adequately implementable*" (A4). Evidence-based guidelines also could lack 'implementability' and therefore be unhelpful. *"Some of the North of England guidelines, for example the asthma guideline, again we had to turn it to a [name of area] one, because it was unusable. ... it was much too difficult to follow*" (A5).

I.2. Change

Change in guideline recommendations without proper justification worried the GPs and evoked scepticism. "*I suppose one does become sceptical with some of the guidelines, particularly with the asthma one, which has been changed in quite large areas. And you then begin to wonder, how long before it changes again and for what reason?*" He added "*... And the blood levels [cholesterol]*

that we are talking about have been reduced and then been reduced again. And it worries me. These are probably knee jerk responses rather than anything else. ...I don't think we prescribe as many statins as other practices do" (P10). Change whether because of new evidence or because of improvement in the guideline might negatively influence implementation. One example was the prevention of CHD: "*That [statins] shows another problem. You know everybody with ischaemic heart disease must have cholesterol under five, which is incredibly tight. And very much moving goalpost and God knows how low can it get again*" (P5).

Noticeable changes in guideline recommendations negatively affected patient-doctor relationship. Patients became sceptical of the quality of care that they had received. One interviewee (P5) explained *"(change) would be very difficult for the patient. What's different today? You need quite a lot of time in explaining. I suppose that's why you go down to special clinics; you've got more time to explain.*" GPs needed more time to educate patients. *"I think the patients would need a lot of education ... we are talking about cholesterols that two or three years ago considered normal being abnormal. Changing your advice to people. ... Not only changing your advice, but putting them on drugs, drugs that could upset the liver. So I think to some extent some of the patients found it puzzling and threatening, possibly*" (P6). GP 'time' was potentially an important factor here. Time is discussed further under organisational factors (4.3.7).

Despite these GPs welcomed guidelines because of 'changes' in the management of patients. "*There tend to be areas where there's change going on. People feel they need update on them. Helicobacter would be a classic for that. People weren't exactly sure what to do with it. ... Some secondary care advice didn't seem appropriate ... And so they [clinical guidelines] had an impact.*" (A5)

If the guidelines had reflected the current state of practice and suggested minor improvements, they would have found their way much easier into practice. "*It's the BTS ones. ... Because it's the only one that I know of. ... To be honest, that's the way we've treated the asthma for years, anyway. I don't think there is anything in it that is any different from how we managed asthma*

for the last fifteen years. So I suppose that didn't make any changes" (P8). Guidelines also helped practitioners to be confident of the care they offered. "*I found them useful, because they confirmed what I was doing was the right thing*" (P10). On the other hand if a guideline had tried to reverse the tide, it would have faced more resistance. Clinical guidelines *"hopefully ... work with the current influence in practice rather than against them.*" The interviewees gave examples of the guidelines that tried to discourage prescribing of selective serotonin reuptake inhibitors (SSRI) or proton-pump inhibitors (PPI), as example of guidelines that tried to reverse the tide and failed.

Ironically, guidelines also failed to be accepted in practice because of the opposite characteristic: reflecting normal practice. Many clinical guidelines repeated the established facts that GPs already knew. "*But I must admit that I do find some of the guidelines that are merely repeating something that been very well established. ... I do find some of them a little bit, just repetitive*" (P10).

I.3. Flexibility

Some interviewees argued that clinical guidelines worked over masses of people and were not designed for individual patients. Therefore it was unreasonable to follow them rigidly. '*You can't be too strict to the guideline. Guidelines correlate over masses of people. That is supposed to be beneficial for the biggest percentage, but it doesn't necessarily help an individual*' (P7). "*Some of the most confusing guidelines are related to diabetes ... They are attempting to say what might be good for thousands of people*" (A4). Research evidence was considered as irrelevant in the same way. "*It could be quite difficult in applying them to individual patients. So whilst they may have validity in population level it's often difficult see how that individual patient sitting in my consulting room fits in. ... The 4S study had a major effect on us but even then we weren't quite sure how it fitted into individual patients*" (P5).

The other argument was that patients' needs were different from those expressed in the guidelines. Guidelines tended to answer straightforward questions. "*It came with no surprise. We do not need a guideline to tell us that if you have more than one or two risk factors and you have raised cholesterol that*

you should consider the use of statins. But that does not help you if you have a 45 year old man in front of you but no risk factors, except that he has a raised cholesterol, the guidelines say nothing about him" (A4). Many patients had co-morbidities, making it more difficult to apply guidelines. One interviewee (P11) explained "*many elderly patients, who the guidelines are particularly for, have more than one problem. Trying to deal with the problems that they want to deal with and working through my agenda of trying to implement the guidelines is difficult and usually the priority in the consultation is the patient's agenda and not mine*". While clinical guidelines expected GPs to think in a certain way, patients' expectations were different. "*If we try too hard with that patient, we will end up using probably 20 tablets a day. And in that patient the top priority might be the fact that he or she has osteoarthritis which is to have a pain killer. While our priority is completely different based on what's we think is good medicine, as based on guidelines*" (A4).

Others argued that doctors were trained to answer patient's needs, not to follow strict rules. Doctors needed to be flexible towards patients' requests. "*But the guidelines can't cope with variation. Black or white. ... They can set a measure for cholesterol level, for example 5. That's OK, thanks. For something like epilepsy: do you have epilepsy or do not have epilepsy. That's OK [but they can't cope with variation in treatments.] ... Anything that we use [a guideline for it] has to be easy and straight to follow. ... We are trained to look and assess people and individual problems for that person and not trained to follow that robust? And make the person fit*' (P4). The guidelines were most useful where the clinical problem had not much variation. In real life practice many clinical conditions were complicated. Unless the guidelines were 'flexible' they would not help management of complicated conditions. "*I think if a guideline is very rigid, it prevents that sort of flexibility that GPs like ... to be able to work with patient rather than just impose on them*" (A3). Some guidelines achieved this level of flexibility. One interviewee gave an example of a guideline produced by GPs for the treatment of helicobacter pylori: '*... and then later GPs' ones came along and they are much more pragmatic and sensible for use in primary care. And so they had an impact. Thinking about that, it's moving from a very strict set of guidelines to a looser more general practice friendly one. That's attractive for GPs*" (A5). "*[If] there are a number of different options for you to use in*

different scenarios [and] *it is not completely rigid [then it is useful]*' (P9). Despite this, one study reported that prescriptive recommendations (such as the NSF for coronary heart disease) were more likely to be followed and were preferred by GPs over loose recommendations (such as the NSF for mental health) (Sheaff et al, 2003).

4.3.4. Theme II: Credibility of source of clinical guidelines

Under this theme the views of the interviewees towards potential sources of clinical guidelines for primary care were discussed (other than GPs themselves). This section assessed the credibility of different professional groups, public and private bodies for development of primary care clinical guidelines.

II.1. Reputable bodies

The source of guidelines was perceived important. "*You got to have trust and faith in the fact that [clinical guidelines] ...are from reputable body*" (P2). It was referred to as one of the reasons for the BTS asthma guidelines' perceived success. Because "*the British Thoracic Society seeing as authoritative without being overly aggressive*" (A5) and "*because the BTS is a reputable organisation*" (P9). The characteristics of 'reputable bodies' identified in the interviews could be summarised as widely known, authoritative and often national, and those bodies that GPs expected to receive advice or support from them. Guidelines from reputable bodies were more likely to be implemented, especially when they possessed other characteristics of guideline with credible source (see below).

II.2. Secondary care as reputable body

Many GPs did not seem to consider hospital clinicians (especially local consultants) as reputable sources for clinical guidelines. They considered high representation of consultants in guideline committees as a factor that might

undermine validity. "*The [NICE's] ischaemic heart one had two or three GPs [in guideline development team] out of 60 or 70 people, which I find deeply depressing*" (P5). On the other hand, referring to the examples of successful guidelines, neither the BTS asthma, nor the Joint British Recommendations were originated from primary care. Some interviewees acknowledged this. Other interviewees had a long list of reasons to justify why they did not consider secondary care as a credible source for useful primary care guidelines:

Patient-doctor relationship

Consultants were thought as unaware of the kind of communication that GPs maintained with their patients. "*A bad example, we were handed a guideline this morning by our midwife, who has some trouble connecting with the practice, saying 'oh here are the guidelines for HIV screening for pregnant women'. ... These are not guidelines with any use to us at all. These are guidelines with two or three consultants who say what's happening. They haven't got the first clue about our communication with our patients. So they are probably not even be looked at*" (P3). GPs were more related to their patients than consultants, and had a less formal relationship with patients. As in this example of HIV screening, it might be reasonable for a consultant to produce a guideline advising GPs to ask sensitive questions, but it would be more difficult for GPs to ask those of pregnant patients.

Case-mix and working conditions

GPs had worked in hospitals, while consultants had not been in general practice. "*Sometimes the hospital clinicians don't properly understand the way we work and the guideline that is simple to them theoretically could be very difficult in the community.... The fact is that all the GPs have worked in the hospital themselves, but the opposite isn't true*" (P6). It was perceived that consultants did not have a clear view of primary care. "*I remember the one a couple of years ago about gastroesophaegeal reflux disease. And it was very secondary care based. ... when you start reading a guideline like that, its advice's based on what's happening in secondary care, you can quickly get bored of it*" (P12). "*Guidelines are less likely to be effective if they come from the bodies that don't reflect the situation where the GPs working. So specialists'*

guidelines are sometime inappropriate" (A2). Also case-mix was different in primary care. "*I would say that guidelines produced by secondary care specialists without input from primary care or public health tend to be pretty useless, because secondary care tends not to appreciate the case-mix that GPs have to deal with*" (A9). Consultants saw 'filtered' patients. "*The bit that they miss out which is the major problem, try to take on board what is actually like in primary care, we see unfiltered patients with their own idiosyncratic health beliefs*" (P5).

One way route of communication and mutual respect

Some interviewees thought that there was a one-way route of communication between secondary care and general practice. "*A lot of hospital departments send regular newsletters out and in that they often contain clinical advice. I found those quite difficult. ... Each department feels because that they have sent a newsletter out we have produced a dialogue, ... I think they are the saddest ones that consultant colleagues try hard to communicate and say how they see the world.... Everybody needs guidelines ... we need to start to respect each others' skills and perspectives more ... This apparently comes to good practice and if you don't adhere to them then by definition your practice is poor*" (P5). This feeling could also result because of 'professional boundaries'. GPs did not produce guidelines for consultants, while they received guidelines from them. GPs also felt that while consultants tried to improve GP behaviour via guidelines, they themselves were in need of improving their behaviour. "*I think they are even worse in following clinical guidelines*" (P9).

Workload

Some secondary care guidelines were perceived as being produced to reduce workload in secondary care by transferring it to general practice.

One stop shops

One interviewee mentioned that hospital consultants had a ladder system for diagnosis. "*The consultants basically create a false situation in that they mostly work by 'trying this and this and if it doesn't work doing something else'; so they tend to have a ladder system for use of guideline. ... GPs do not work that way.*

... GPs are more like a one-stop shop. ... In fact keep out the consultants. Don't let them chair the meetings, only use them as resource" (A8). The interviewee fell short of saying what GPs did when their patients came back for further advice for the same condition. '*Keep out the consultants*' was amongst the extreme views and was not shared by everyone.

Restricted pharmacy

"*They have a restricted pharmacy. They can't prescribe everything. And what they can prescribe is often dictated by cost. And we know that the hospitals are able to negotiate very favourable prices for drugs, which is not available in the community when we start prescribing it*" (P8). Therefore, a drug which had a favourable price for hospital prescribing could evoke unnecessary costs in general practice. This situation, however, has now more or less changed. With the increasing power of PCTs and also the tendency of hospital Trust and PCTs to issue joint formularies, GPs and consultants tend to have similar choices.

II.3. National bodies as credible sources

At the time that the qualitative study started NICE was a young organisation and had not yet published any 'guidelines', but 'guidance'. The interviewees referred to those NICE publications as clinical guidelines. No direct questions were asked about NICE or other national bodies in the interviews, but the interviewees referred to them frequently.

National bodies, including NICE and also Scottish Intercollegiate Guidelines Network (SIGN) were perceived as credible sources. There were some positive comments on the quality of SIGN guidelines. "*There are Scottish guidelines called SIGN. Because they are on the web, while I don't use them, I refer to them quite a lot. And they extremely good and I think they are probably an under-used resource in England*" (A3). NICE was perceived as a national authoritative body, which could ease the problem of multiple sources of information. "*We get an awful lot of information and it's difficult to pull out what's the most important piece. That's why actually the NICE guidelines are in a better way of doing it. Because they come from a central recognised [body],*

whether it's approved by everybody, but it's recognised by everybody, and they come in the same format with the blue cover and it's thin ..." (P8).

There was a sense that some National Service Frameworks (NSFs) and NICE guidance had been imposed on general practice without proper justification. They saw them as influential sources of change in prescribing. "*I suspect we will be driven more and more by guidelines that come out of NICE, because there is no option but to listen to those whatever you think of them*" (P5). Interviewees were also unsure about the way the NICE guidance had been produced or how the topics were being chosen. They were particularly negative towards NICE guidance for Relenza prescribing. "*And it's just really unrealistic of people at NICE expecting people follow that sort of guidelines [Relenza guidance]*" (A9). One academic GP had a positive view on Relenza guidance, as "*something that GPs can stick to*" (A2).

II.4. Pharmaceutical companies

Pharmaceutical industry was not perceived credible for producing clinical guidelines. "*Examples of useless guidelines are often driven by pharmaceutical companies*" (P5). Even the industry's contribution to guidelines developed by others resulted in scepticism and could jeopardise the credibility of source. Some interviewees thought that possible contribution of pharmaceutical companies to the production of BTS asthma guidelines damaged their credibility. "*They are largely funded by Glaxco and there are issues about, you know, their emphasis about drug treatment. But that's the world that we live in and that's where the trial evidence largely exists*" (A12). Drug companies produced lots of different guidelines, which were generally perceived as unhelpful. This 'mass production' could be a 'noise' factor in guideline implementation, as they might distract GPs attention from useful and evidence-based guidelines.

II.5. Publishing guidelines

The most prominent of all respected sources was the British National Formulary (BNF). GPs tended to consult the BNF frequently in their practices. They found

it helpful that asthma and CHD prevention guidelines were included in the BNF. "*The fact that [risk tables] are now being printed in the BNF, they are always on my desk*" (P11). "*[BTS guidelines] are actually incorporated into the BNF, which I find very useful. ... I think the ones that are sort of really ratified then they should all go into the BNF*" (A11).

Generally speaking, publication of clinical guidelines in respected journals was considered as a source of credibility. Among those respected sources, the BMJ was eminent, probably because it was more widely read. "*A shorter version of [BTS published] by the BMJ ... [we followed it] because it comes with the weight of an organisation like the BMJ behind it. And I think the BMJ tends to have a good reputation in a sense that it tends to be fairly peer reviewed*" (A9). There were also references to the Drugs and Therapeutic Bulletin, the Lancet and Thorax. "*We discussed benzodiazepines in the late 80s when the Lancet article came out saying that they are dangerous if [prescribed for] more than 10 to 15 days*" (P3). But when it came to other journals it was less clear. Publication in respected sources was not adequate for implementation. *"[North of England guidelines] were published in the BMJ, [GP] sees its article and takes it to his or her practice. And decides after discussion with the partners they were going to do something within the practice. [It is not enough]*" (P2).

Although 'Clinical Evidence' (BMJ Publishing Group) was a new publication at the time of the interviews, there were positive references to it: "*Another book that I think is wonderful is the recent BMA publication about clinical evidence*" (A11). Clinical Evidence could be a useful source whenever single and short evidence-based message was required.

4.3.5. Theme III: Presentation of clinical guidelines

III.1. Simplicity

"*The message has to be simple. If it is overly complicated, people tend to ignore it*" (A5). There was general agreement that all guidelines should be simple to

follow. "*The guidelines are quite difficult to use. You never find them when you want, and when you do find them they are often badly designed and you can't actually fit a particular recommendation to a particular patient*" (A12). Strong and simple messages were more likely to be considered. One interviewee explained: "*... I'm sorry if you show me a complicated piece of paper it's no use to me. I'm a simple man and I need to have simple ideas*" (P5). It might seem a bit odd for an outsider to see GPs insisted on the need for simplicity of clinical guideline recommendations. The need for simplicity might reveal the complexity of the environment in which GPs practiced. Complexity of the message might also hinder understanding and hence make the message less persuasive (Raven and Rubin, 1983, pp185-187). GPs' emphasis on simplicity might also reflect short consultation time, which is often less than ten minutes (Carr-Hill et al, 1998). Simplicity of presentation was among the features of NICE publications that GPs welcomed. "*[NICE guidance's] come in the same format with the blue cover and it's thin. There is an executive summary at the beginning and there is evidence for the reasoning. I've been quite impressed by the way we all get them and they are available by emails as well*" (P8).

Simplicity was not a simple concept. There was not complete agreement on examples. Guidelines seen as simple by some interviewees were considered as difficult by others. Many interviewees considered the BTS asthma as simple, but two interviewees (P5 and A4) thought otherwise. "*[They] are actually quite complicated and difficult to follow, and there are couples of drugs in there where their position is not really very clear...*" (P5).

III.2. Systematic presentation

Asthma treatment recommendations were in stepped approach in the BTS guidelines. This systematic presentation was supported by the interviewees. The interviewees expressed that the stepped approach helped 'logical thinking'. "*It does have some influence... where I am and what stage I am. It makes it more logical in the process of treating the patients*" (P10). Guidelines for CHD prevention were also thought to have organised presentation: "*I think that guidelines are good ways of codifying, making explicit evidence around the effectiveness of statins*" (A12).

The innovative stepped approach in the BTS asthma guidelines was a successful strategy to overcome the complexity of asthma classification and treatment. Stepped approach could be compared with splitting asthma to a few sub-categories within which patients were located in various stages of the disease. So instead of having a complicated clinical scenario with a variety of treatment options (difficult), there were a few sub-diagnoses with small number of alternative treatments for each (simple). Once it was decided where to put the patient, choice of treatment was not difficult. This was mentioned as the main reason why the BTS asthma guidelines were perceived simple: "*The structure of them makes them clinically easy to use*" (A11). There were no known studies assessing the effects of 'stepped presentation' of the BTS asthma guidelines on their implementation. However, experimental research in nursing demonstrated that the way patient information was presented to the nurses influenced their behaviour and quality of care they offered to the patients (Dowding, 2001).

4.3.6. Theme IV: Influential people in implementation

According to the data, different groups of people influenced GPs' intentions and plans to follow specific guidelines and their decisions to prescribe medicines. These reference groups had positive or negative influences in different circumstances. Six main reference groups were identified and discussed under this theme: 'patients', 'consultants', 'GP colleagues', 'nurses and other members of primary care team', 'primary care organisations, pharmacists and prescribing advisers' and 'pharmaceutical companies and their representatives'.

IV.1. Patients

Patients who did not comply with treatments were perceived as barriers to implementation. Different factors were involved, but patients' difficulties in understanding clinical recommendations were perceived as major factor. "*[Barrier to] asthma [guideline implementation] is patient understanding; biggest*

problem and it takes us time" (P6). "*Patients very often struggle to understand the [CHD] risk … an awful lot of patients don't have a sufficient risk to justify the primary prevention use of statins. But patients perceive that slightly raised cholesterol wants a tablet to deal with that*" (P11). Also some "*patients do not like taking prophylaxis*" (A12) for asthma or for CHD. "*My barrier for prescribing statins is that people don't like taking them. … So I think as a clinician, I'm less inclined to prescribe them than the guidelines make me*" (A11). All these resulted in prescribing not in accordance with guidelines. Patients influenced GPs' choice of drug preparation. "*Another guideline we got from prescribing authority, to switch off onto generic preparations for salbutamol … and you see patient comes back and says 'I've used it for years and it works better'. Maybe it is? Maybe there is something in the way that company makes it …*" (P7). In occasions, adherence to guidelines resulted in patients leaving the practice. "*We used to use methadone more, until a guideline came out and we decreased it … lots of our drug-addicts left…*" (P7).

The interviewees noted that as patients became more aware of treatment options (e.g. for CHD prevention) they asked their GPs to adhere to them: "*should I actually be taken something?*" (A6). Some GPs were positive towards this: "*I would be very happy for every patient with a chronic disease to have a copy of an up-to-date guideline. Patient may say doctor why am I not having ACE inhibitor and the doctor would have to explain it*" (A3). Particularly in cases of "*more obscure conditions, patient take charge of their care in the way that obviously encourage us to share that understanding of the guideline*" (P12). The findings implied that improving patient access to guidelines increased patient expectations from doctors. It was known that this sort of pressure from patients was more likely to be exerted by middle class and educated patients (Horder et al, 1986)

Guidelines influenced relationship between clinicians and patients in different directions. Guidelines helped doctors back their decisions and convince reluctant patients: "*it makes more sense to say: this is recommended with some authoritative basis than say this is what I do*" (A3). This might be another reason why guidelines recommending no change in practise and

confirming current prescribing could be beneficial. Guidelines could overcome patients' reluctance in complying with treatment.

IV.2. Consultants

Local consultants were generally perceived as influential on prescribing. "*I actually think that the doctors that GPs refer to, influence their prescribing quite a lot*" (A11). This influence was exerted through letters sent to GPs and educational meetings. "*The key player in all these is the consultant. The consultants are whom that GP sends patients to for advice and patient comes back on prescription...*" (A12). "*And in lunch time meetings and evening meetings... without meetings, the guideline wouldn't be useful*" (P7). Consultants' influence was also dependent on local relationship: "*If the GPs have a good relationship with the local consultants, then I think they are in a fairly very powerful position in implementing guidelines*" (P2). Some interviewees thought that this influence was becoming less prominent: "*I think their influence is much less now than before*" (A4). Consultants had limited influence when GPs managed patients independently. "*Most of our asthma patients are not supervised by the hospital. So I don't think anything from the hospital colleagues*" (P8).

Consultants' influence on prescribing was perceived as not necessarily in accordance with clinical guidelines. GPs thought that consultants seldom recommended clinical guidelines in their letters. "*They very seldom invoke guidelines. I think if they did they would be very helpful. Actually that suggests to me that specialists don't follow guidelines either*" (A3). This was echoed in others' views: "*I think they are even worse in following clinical guidelines*" (P9). Therefore, consultants could potentially hinder guideline implementation in primary care: "*as a GP, as an individual it is very hard to resist that drive*" (P5). "*If you work closely with somebody in secondary care who has a more idiosyncratic method of treatment, you may end up prescribing the way they do instead of using the guidelines*" (A11).

The findings should be viewed in conjunction with what presented under the sub-theme 'secondary care as reputable body' (4.3.4). GPs did not support

secondary care initiated clinical guidelines, and were less willing to adhere to those guidelines. It was also said that GPs felt that consultants did not understand primary care. Here it was noted that GPs believed consultants influenced their prescribing. These seemingly conflicting positions were very important in understanding the dynamics of primary care. Decision makers should be interested in the possibility reconciling these two positions to improve the quality of care. Would consultants be co-operative in the implementation of guidelines that were produced without their input?

IV.3. GP colleagues

GP colleagues within practice were considered influential. The influence was variable, and as for consultants was not always in line with guideline recommendations. "*Partners are really important. ... If your partners prescribe differently you might pull the guideline out and say that's actually what this guideline says. Unless you have the confidence to do that, you might be more inclined to what the practice wanted*" (A11). GPs used their in-house meetings to reach agreement over specific prescribing approaches or whether to follow a guideline. "*We have a little clinical meeting every couple of weeks and try both [my GP colleague and I] doing roughly the same thing*" (P7). Meetings could happen as often as every day (P3), in informal way, or be planned as a formal session: "*One of the partners did some homework. There was a meeting of all clinicians, it was presented and discussed. And then the written version is shared around and people have it on their desks*" (A3). The meetings reflected the reality of modern general practice, where GP partners 'shared' patient care responsibilities. "*We found that anything is only useful if all three partners agree. ... There is no point at all if I stop prescribing benzodiazepines, unless we all do so*" (P3). "*We all do mostly the same to be honest. And we do discuss it at intervals*" (P8). In this team approach some GPs assigned the responsibility of looking for evidence to one partner. "*We nominate one of the partners in the group to have a look about question specifically and she is being much more aware of the guidelines [for statins]. We work as a practice team on that level and we distribute audit and guideline update*" (P12). "*... One of partners is interested in contraception ... so she keeps us updated there*" (P3). In-house meetings did not happen regularly in some practices (A3).

GPs found it easier to reach agreement about a clinical guideline in smaller partnerships. "*Probably the bigger the group are in, the more disagreement you'll get. ... Can you thrash those out or do you accept lowest common denominator rubbish*" (P5)? Another interviewee who was from a large practice did not see in-house colleagues as influential: "*If you have to find someone [influential], I wouldn't think it would be anybody necessarily in the practice*" (P2). Individual GPs might play dominant roles in their practices: "*one of partners probably has been a little bit more [influence]*" (P10).

The influence from GP colleagues outside the practice was fairly limited, unless from motivated colleagues. "*One of our colleagues in practice nearby does quite a lot of research on epilepsy. He often talks about epilepsy*" (P6). Influence of GPs outside the practice seemed to be occasional: "*if we talked to our colleagues in other practices and see something complete different we missed it, then we like to catch up*" (P3).

IV.4. Practice nurses and primary care team

Practice nurses played important roles in guideline implementation, specifically where they were actively involved in patient care and running mini-clinics (e.g. asthma, diabetes and CHD). Some GPs found this encouraging: "*I think the important input that we've found is really from our nursing colleagues*" (A13). Nurses also provided new evidence and useful information. For example "*... when there is a change for diabetes –again WHO criteria changed- or for asthma that makes almost immediate impact. Partly because our nurses, who share that care with us, bully us if we haven't noticed*" (P3).

The interviewees described nurses as "*extremely good in creating structured approach*" (A6) and as those who "*follow protocols much better than doctors do*" (A3). "*My guess is that the vast majority of prescriptions for [CHD prevention] the nurse says to GP I need this. GP doesn't think about it, they just sign it*" (A10). For that reason some guidelines specifically targeted practice nurses. A13 provided examples of such guidelines for epilepsy care. GPs felt

confident in transferring some of their responsibilities in chronic disease care to nurses. "*We trained our nurses to look after our asthma patients*" (P11).

One interviewee mentioned that potential disagreements between GPs and nurses negatively affected guideline implementation. "*There are some professional boundaries about. Who is going to do what and whether we all agree about it.... Traditionally the doctors; but nurses seem to get more involved in statins, nurses tend to be running the CHD clinics. [On the other hand,] GP might accept the guideline, because he wants to give it to the nurse to run the clinic with. But obviously that involves both sides to agree*" (A5).

Other primary care team members were mentioned in three interviews. The importance of a team approach was emphasized for successful implementation CHD prevention. "*[Statins] is a fabulous example of everybody in primary care playing a part and that's the success of the implementation of the guideline*" (A6). "*Statins for example, you've got to have all the practice on board, practice manager doing things and so on*"(A5). District nurses also influenced implementation (P11).

The interviewees who perceived nurses as influential in guideline implementation considered the influence to be positive. This was different from the perceived influence of consultants and GP colleagues, which might or might not be in line with guideline recommendations. On the other hand, about two-third of the interviewees did not refer to nurses, whether positive or otherwise.

IV.5. PCOs, pharmacists and prescribing advisers

The term Primary care organisation (PCO) was used to refer to Primary Care Groups and Trusts in England, Local Health Groups or Boards in Wales and Local Health Core Cooperatives in Scotland.

Interviewees thought that PCOs had limited influence on prescribing. The influence was materialised through recommending specific types of medicine and discouraging others. "*Our local PCG has been issuing guidelines on which particular type of drugs to use in asthma*" (A4). PCO formularies were not

thought as essentially helpful. "*We have just one copy of that formulary in our practice and I don't know if they are going to put in every doctor room and expect people to stick to it or not. I suspect not. Because I can't imagine a lot of doctors say 'Right. You need anti-hypertensive, let me see what my PCG let me prescribe*" (P1). "*A problem with a PCG formulary is that lots of practices over a wide area are asked to do lots of things together. ... there are difficulties involved in trying to get everybody to feel as if they have a part of the formulary or any decision making is very difficult*" (P2).

PCOs' prescribing messages were normally received via their prescribing advisers and attached pharmacists. Advisers and pharmacists provided GPs with information and guidelines. Some interviewees felt too much information was delivered in this way: "*a lot of guidelines from prescribing advisers come around*" (P7). One interviewee mentioned that the PCOs' influence was growing: "*of much growing influence are the prescribing managers in PCGs and the attached pharmacists of the practices*" (A5). "*In Wales ... generally we have employed a pharmacist who worked with the practice. So someone like that I think could have an influence*" (A8).

One interviewee expressed very negative views on PCOs' role in general. "*... I don't think that we can honestly say that PCGs have the time to influence implementation of the guidelines in any great way at the moment. ...I think it's waste of time, mainly because they are under-resourced, and the goals of Department of Health are unachievable. We end up subsidising PCGs with our good will and time and money trying to implement things that are difficult to achieve at the best of times; and with the very limited resources that we have, becomes impossible. So I don't think that PCGs are here to stay*" (P2).

IV.6. Drug companies and reps

Drug companies, mainly through their representatives, were another source of influence. "*The other influence on prescribing that you have to consider is drug representatives. We do have influence from them. And it would be naïve for me to say they don't. They do influence us*" (P6). Some GPs did not see drug company representatives. There was an effect on GPs who saw

representatives. The interviewees perceived that the influence was limited to certain areas where it coincided with other factors. Two main examples were mentioned. Drug companies' campaign a few years ago to encourage GPs to follow BTS asthma guidelines was one of example. The other more recent example was the industry's interest in GPs' adherence to primary prevention charts, as they believed it increased drug prescribing and hence their profit. "*Those charts [of CHD] became very familiar partly because pharmaceutical companies started giving free copies of them away*" (A5). This sort of influence was perceived as informative by one interviewee. Another interviewee mentioned, "*their influence I think increasingly is based on scientific facts*" (A4).

Some GPs thought that representatives had an effect by targeting consultants (A1) or nurses. "*The other key fact was the industry, promoted [BTS] heavily, particularly to practice nurses. So you could get free copy of the guideline ... because they said use more medications*" (A5).

4.3.7. Theme V: Organisational factors

This theme covered organisational factors that facilitated or prohibited adherence to guideline prescribing recommendations. Five sub-themes were identified in this category (Table 4.1).

V.1. Practice characteristics

There was no reference to dispensing status or previous fund-holding status as factors inhibited or facilitated adherence to guidelines. One GP trainer considered training status as a positive factor: "*we are a training practice; we try to keep on top of things*" (P10).

Special mini-clinics for asthma and CHD prevention helped guideline implementation. "*We now run special ischaemic heart disease clinics ... so we are very hot on statin because of the accumulating evidence*" (P5). Mini-clinics helped clinicians (including practice nurses) to focus on the disease and provide

a more coherent service to patients. "*I mean becoming more formal in our dealing with asthma patients using our asthma clinic, and therefore probably using more preventives, steroids etc. on a more regular basis than we were previously*" (P10).

V.2. Information technology

The interviewees were positive about the usage of information technology for guideline implementation; perhaps more as a way of the future. Fruitful IT experiences were limited: risk calculation for primary prevention ("*on a regular basis I use a piece of software on the computer which is a risk calculator*" (A9)), searching for a guideline when you need it and some limited use of PRODIGY (Prescribing Rationally with Decision support in General Practice) software package. They saw computerised systems, which included clinical guidelines, as the way forward. "*Obviously with things such as PRODIGY being around, longer term acceptability of the guidelines will be influenced by whether or not they are usable within a computerised environment*" (A5). The other important usage of computers was through using computerised patient records for identification and proactive follow up of patients, especially those with chronic or rare diseases.

PRODIGY was perceived as useful software package for guideline implementation, mainly for diseases with straightforward diagnoses and limited choices of treatment. "*I used [PRODIGY] twice today. Once for scabies and once for tonsillitis and in both cases the computer guideline offered the recommended treatment. In both cases there isn't any choice. ... and the computer guideline actually made it easier than think about it myself. ... In more complicated cases like hypertension it goes to a rather long branching tree. So it takes longer rather than just writing your preferred drug*" (A3). The perceived inefficiency of the computerised packages for enhancing the implementation of guidelines for chronic diseases had also been reported in a randomised trial (Hetlevik et al, 2000). GPs sometimes found PRODIGY intrusive during their consultation sessions. "*... even PRODIGY, I found it quite intrusive to be honest. I found it very useful if you got a difficult patient and, I think, after the surgery. But actually during consultation do I want prescribing prompts? The*

answer is yes, but if I have got the time to go through it" (P5). As a result many GPs chose not to follow PRODIGY instructions.

V.3. Availability of required resources

Implementation of some guidelines relied upon specific instruments and hardware. The interviewees offered some examples: COPD guidelines (spirometer), guidelines for heart failure (echocardiography) and the diagnosis of urinary tract infection in children (collection bags). "*I found COPD guideline by BTS quite difficult because it requires spirometery, and being realistic, many of us don't have spirometers. I'm not sure how relevant they are to routine care...*" (P6). "*In heart failure ... you can make accurate diagnosis only by echocardiography and many GPs realise that the appropriate treatment is ACE inhibitors, but GPs don't always have access to the investigations and there is a long waiting list. ... Whether it's the guideline that's failing or the implementation process itself was handicapped*" (A4). "*Heart failure [guideline] is a wonderful guideline, but how do you actually put it into practice if you can't get open access to echocardiography*" (A6). Similar problems happened in the absence of human resources. "*I think quite a lot of barriers are organisational: in order to meet the guideline, you might have to have a system of care*" (A6).

V.4. Time, workload and information overload

There was a genuine belief that general practice was under increasing pressure, so that it made it difficult to plan for quality improvement in general and guideline implementation in particular. "*You know a barrier to all clinicians doing things better is that they don't have the time*" (A12). "*At the moment we've been bombarded with so much change that our practice manager, nobody, can manage. So if there are any guidelines to come, that's not a good time for them...*" (P3). "*Time is issue. It's impractical to do that, without losing for other things. ... I think modern general practice requires you to be proactive, to be planning for staff, get involved in staff training, following guidelines ... safety of work .... They take quite a lot of time. We have a meeting regularly which is purely non-clinical thing. ... So that influences my prescribing, because I don't*

manage to prescribe very much because I'm always at meetings (laughing)" (P5).

The process of reading, understanding, discussing and utilising guidelines and eventually auditing implementation could be time consuming. Adding to this, the consultation times were pretty short. "*The biggest [barrier] is the lack of time to initially study the guidelines and take them in, and then when the patients come in the lack of time to work trough the guidelines*" (P11). The interviewees considered the majority of guidelines as technologies that produced extra workload. "*Time to read, understand, to design strategies, to do audit ... is enormously difficult and the resources that go along with that ... workload in general practice is so enormous that to try and remember to think about other things is really tough*" (A6). When there was a sense that following a guideline decreased demand on GP time, guidelines were more likely to be followed. "*In helicobacter, there are some suggestions that guidelines might decrease workload from dyspepsia and so on. Whether that was true or not, is another matter. But there was a perception that it might reduce demand, so it was easier to uptake*" (A5).

The step-by-step approach towards treatment (i.e. first treat by the least expensive drug – which probably was less potent with more side-effects) recommended by many guidelines was considered a barrier to implementation. GPs thought this approach resulted in more patient appointments and pressure on one of the most precious items in general practice, i.e. time. "*If you look at most treatment drug guidelines they actually go down a pattern of try this first then do that. And that creates more appointments*" (A8). GPs' concerns over their workload and time constraints were well known. But time did not seem to be the crucial point. If there was perceived need for a guideline, then time would not be a major barrier. GPs worked 'smarter' (Poplin, 2000), shifted or changed priorities to produce necessary time.

Information overload or simply guideline overload was another matter. Many guidelines arrived through the post and were ignored. Some GPs sent some guidelines "*straight to the bin*" (P5). Others tried to keep them, but failed to use them: "*you can't have a guideline library*" (P7). "*Having a memorable*

article in the BMJ ... months ago where a GP was photographed beside the pile of guidelines that had arrived by post during the previous year and it was about 4 feet tall" (A2). Given that GPs received many guidelines, it was more difficult to choose the suitable ones out of the others. GPs tried to be selective, but there were not explicit criteria. "*That's the main problem in whole business ... we have to select one and ignore the rest unless there seems to be a reason for change. ... There are lots of guidelines. Protocol for that, guideline for this... Information overload, guideline overload*" (P4). In some clinical areas the number of produced guidelines was high, and it was perceived unreasonable to expect GPs to consider new guidelines. "*If ... yet another antibiotic guideline, frankly not even be looked at, because it's been reviewed so many times*" (P3).

The idea of a recognised body taking charge and introducing reliable clinical guidelines was appealing to some interviewees. "*I really like it if some overarching body will pull together the best guidelines, all the ones that have been really quite marked and say give that book to GP updated every year*" (A11). In this sense, GPs had a positive approach towards the NICE role.

V.5. Costs and expenditure

Especially in the case of statins, drug expenditure was an important issue. GPs saw themselves under scrutiny to prescribe within budgetary limits. "*And yet [while asked to prescribe statins] GPs are criticised for over spending on drug budget*" (P5). "*You can't do something like [CHD prevention] without spending more money at the end of the day*" (P6). GPs were more likely to follow guidelines only for secondary prevention of CHD, and not primary prevention. "*I don't think the [guidelines] help. Because we can't afford the use of statins that the guidelines suggest ...*" (A8). "*I think there is a real concern about the cost of statins and how that might influence prescribing budgets and GPs ability to prescribe other drugs if they would implement the guidelines as widely as they would wish to do. I think the evidence so far is that if people stick to secondary prevention then it shouldn't be something that actually expensive*" (A13).

The issue of cost was more prominent when guidelines requested initiation of new and expensive drugs without removing the need for other

treatments. "*Biggest barrier there [statins] for implementation is initially drug cost. ... It was a new area. Wasn't replacing one drug with another. Could argue in asthma you are doing it; if you implement the guideline ... reduce the use of beta2*" (P6). PCOs were not perceived as supportive of guideline implementation. "*[PCO] where most of the funds come from says you can't afford to use [statins]. So in fact in a way they are counterproductive. They get you thinking down statins guidelines, but they don't help you*" (A8). While many saw adherence to CHD guidelines as costly, some considered guidelines as means of limiting prescribing costs. "*There is a perception that the statins guideline is based on making sure that not too much money spent on statins*" (P11).

Some interviewees argued that if clinical guidelines recommended cheaper drugs, that were as effective as their counterparts, it would act as an incentive. "*If the guideline says a drug is only marginally more effective ... but is dramatically cheaper, then I think that would influence them as well*" (A11). Equally guidelines were welcomed if they justified additional prescribing cost, e.g. in case of asthma clinical guidelines.

4.3.8. Theme VI: Disease characteristics

This theme focused on items which directly explained the specific effects that disease (or clinical condition) characteristics had on implementation. Three sub-themes were identified under this category: 'treatment is secondary care based', 'difficulty of diagnosis' and 'rare or simple disease'.

VI.1. Treatment is secondary care based

GPs had limited roles in the management of certain diseases. Treatment of some diseases was solely or largely secondary care based. Secondary care was the natural initiator of the treatment for other diseases and GPs' role was mainly to follow up and deal with adverse events, exacerbations etc. Epilepsy was a good example of a disease with these characteristics. Interviewees

believed that there was less room for manoeuvre in primary care for adherence to guidelines' prescribing recommendations. "*GPs initiate epilepsy treatment very rarely. Usually secondary care doctors who initiate the treatment and we just re-prescribe. ... I as a GP have rarely requested routine monitoring, so I am not aware of what the guideline says*" (P11). "*Since initiation of anticonvulsant [for epilepsy] treatment is not a primary care task we ... seek advice from specialist on that and we follow their advice*" (A3). GPs were perceived to have important roles in identification of epilepsy patients on multi-drug treatment (which was discouraged) using practice data. As an exception, one of the interviewees had special interest in epilepsy (using SIGN guideline: Scottish Intercollegiate Guidelines Network, 1997a; Scottish Intercollegiate Guidelines Network, 1997b) and seemed to have managed to influence GP colleagues outside his practice.

VI.2. Difficulty of diagnosis

Where the diagnosis was problematic, clinical guidelines with emphasis on treatment (including prescribing) might not make much change. The prime example was depression, as one of the interviewees explained: "*as a student and junior doctor I had a very clear idea about the difference between depression and sadness. I don't think there is a distinction anymore*" (P5). Some of the concerns that were expressed on NICE's guidance for prescription of Relenza, referred to the same problem as they saw influenza as a difficult disease to diagnose in first place.

VI.3. Rare or 'simple' diseases

For some clinical conditions a short piece of evidence and recommendation might work better than a clinical guideline. Menorrhaegia was referred to as an example. While some interviewees were more or less aware of the main recommendations of the guidelines, no particular guideline was recognised as the source. "*I understand that the evidence is that tranexamic acid is the most effective drug and ... And majority of GPs and many many gynaecologists do use noretisterone, which is ineffective. So it's not quite a guideline, much more being aware of evidence*" (A3). Some GPs used the guidelines when they

encountered a problem. "*I think that probably stops GPs prescribing so much progestogen. ... I remember it jumps from treatment with things like mefenamic acid ... to a more severe treatment, which is less well tolerated in primary care. ... I know what the first step is, so I always use that, if I pass the first step I do look back at the guideline*" (A11). For clinical conditions like menorrhagia, short compiled evidence could be more useful. Publications like the 'Clinical Evidence' could provide a proper answer to this need. The other useful approach was more active use of IT in practice as accessible source of information whenever required. IT was particularly useful when doctors sought guidelines for treatment of a rare disease. Also patients could play a role in the provision of guidelines for obscure conditions (4.3.6). The effects of disease complexity on implementation had theoretical and empirical backing in the diffusions of innovation theory (3.2.3).

4.3.9. Theme VII: Implementation strategies

VII.1. Planning implementation

Lack of implementation programmes was regarded as the prime reason for non-adherence to guidelines. "*I think probably the weak step in all guidelines is implementation. The people think that once they've written the guideline and come out with a pretty algorithm, that's it and all they have to do is post it out, and the truth would do the rest*" (P5). "*They got to be effectively disseminated. Generally they are not. They're just posted out*" (A7). Lack of implementation programme was a barrier regardless of the quality of the guideline. "*I don't think North of England [asthma] guidelines had much impact in general practice. Few people knew them*" (A5). "*All guidelines can be useless if they are not read or used, and there is a problem of fact that we getting better and better making wonderfully streamlined very scientific guidelines, but unless they are actually read and digested and put into practice, then they all be useless. And I think some have ... fallen on stony ground*" (A6). "*You can't do sit down with a copy of the guidelines with a covering note saying please do this. I don't think it works*" (A13). There were many examples of weak implementation strategies:

"[depression] is a topic that is a bit of a Cinderella. ... I think it is an example of where there is evidence but not a great implementation strategy and a lot of variation in practice" (A6).

Good dissemination and marketing strategy could prove successful even with an acceptable and not necessarily evidence-based guideline. The best example was the earlier versions of BTS asthma guidelines. *"Because [BTS] is the one that was marketed the best, was advertised the best, it was the one that used nationally. [Even if not evidence-based] but they are workable documents that everybody uses. To use a different one, seems a bit [odd]"* (P3). Continuous multi-faceted implementation programmes were thought to be the most effective. *"[BTS] had a major impact because of its marketing strategy. That was presented in a very simple easy format, widely accessible, widely talked about, the use of multi-factor approach towards implementing the recommendations ..."* (A6).

VII.2. Can implementation be successful?

Thinking *"that the guidelines on their own would make a difference is probably naïve"* (A12). Even when the change in prescribing was as guideline recommended, it was not clear whether the change was due to the guideline, because *"guidelines are only part of this"* (P2). Interviewees gave a few examples of change in their prescribing coinciding with clinical guidelines while they were not sure what factors were the main causes of change: *"I suppose I tend to use aspirin in people with ischaemic heart disease ... but can I link it to the guideline?"* (P8). *"I think one of the problems is how to define implementation of the guidelines. I suspect that over time, for example, general practice prescribing of anti-asthmatics has moved more closely to BTS guideline. But I'd find it very hard to say with any sort of scientific certainty that the BTS guidelines did that. ... they may over time actually influence the practice, but it's hard to have a cause and effects relationship, unless within scientific setting of a research study"* (A10). *"I only believe trials, there is so much confounding out there. I mean, I give you an example, look what happened to statins prescribing. I don't believe for a moment that statins prescribing has gone up because of the Joint British guidelines on management*

of coronary heart disease, although if you look at the graph in a time series way, ... it is really begins going up in 1999 when those guidelines came out. But I think one has to be very careful before one can say a specific guideline has actually made the difference" (A12).

Some interviewees thought that they had altered their prescribing as a result of guideline recommendations: "*definitely with asthma I'm moved toward preventative treatment*" (P6). Generally speaking the views of non-academic GPs were more positive towards the influence of clinical guidelines on prescribing, while academics tended to be less confident of guidelines as the cause of change in prescribing.

There were specific cases where guideline effect was prominent. NICE's guidance on Relenza was one example, as it was a relatively new line of treatment. Before the guidance was released, many GPs were reluctant to use Relenza on the recommended scale and with its release the drug was prescribed more. Therefore, interpreting the cause of change could be context dependent. For menorrhagia the guidelines were not actively implemented and were unlikely to have caused notable changes in practice. For statins, although the guidelines were actively disseminated, perhaps accumulative evidence and different sources of message had caused most of the effects.

In some cases GPs were aware of guideline but not familiar with the message. More commonly they were familiar with the guidelines' message without remembering the source, specifically where there were many different guidelines for a condition, or where there had not been massive implementation strategies. This could hinder the implementation of guideline updates.

Guidelines could play another role in implementation. They might help 'change' to stay longer in practice. They could provide a milestone for reference and implementation. Respected and well-known guidelines (e.g. the BTS asthma) transferred new evidence, through update versions, much more easily into practice.

Guideline implementation could go wrong in different directions, for example, when the message or dissemination approach were prone to misunderstanding. It might cause firm beliefs against the wishes of implementers and the content of recommendations. These beliefs could be more resistant to change. One interviewee reported that some GPs used CHD primary prevention risk tables for patients who fell within secondary prevention category. A previous study reported similar phenomenon (Fairhurst and Huby, 1998). When studying cost-effectiveness of clinical guidelines, these are among the adverse events that should be measured and taken into account.

VII.3. When do GPs welcome new guidelines? Perceived need, past experience and knowledge

GPs were positive towards change coming through guidelines as it helped them to perform better in their professional responsibilities. In this sense they were receptive towards guidelines. "*We as professionals want to know we are doing our best for our patients ...*" (P2). Guidelines were more effective when GPs had perceived need for advice for treating a disease, or when they were aware of potential shortcoming in the care they provided. "*Somebody who seeks guidelines is much more likely to implement them than if guidelines imposed*" (P2). This 'need effect' could be the result of another phenomenon. As noted because of information overload GPs ignored many guidelines and chose only those that considered helpful (4.3.7). When there was a need, GPs were less likely to ignore a source of evidence-based advice even if it came through the post. "*I was ready for [BTS asthma guideline]. By the time they came out I was desperate to read them*" (P5). "*I look at the ones that I have been really looking for*" (P12).

Targeting clinical areas where quality of care was lower than expected could result in dramatic achievements and boost general practice morale. One interviewee (A13) reported 90 percent change (improvement) in prescribing when they introduced a guideline for the use of sublingual analgesics. He believed this success was due to good communication with GPs and GPs readiness for adopting the guideline. "*Why do you introduce guidelines and usually we produced a guideline to resolve a problem. And if that problem have*

been recognised by the GPs they have usually been co-operating with the guideline that we produced" (A13). To complete this discussion, it should be noted that perceived need was not a pre-requisite for every implementation. As for the example of NICE guidance on Relenza, while there was no particular perceived need in primary care, the guidance was to some degree adopted.

GPs were 'selective' in using guidelines. They did not read every guideline they received. Many guidelines were thrown away after a quick look. "*I usually have a look at [clinical guidelines] and if I don't think they're gonna to be useful I [throw] them in the bin*" (A11). If a guideline was read once and perceived as non-helpful, more effort would be required to convince GPs to read it again. We called this the 'first contact' effect. Clinical guidelines should be presented and disseminated in a way that they attracted GPs' attention. There was theoretical and experimental support for the importance of the first contact. In social psychology it had been referred to as 'primacy effect' (Raven and Rubin, 1983, pp 90-91).

Guidelines were also used as a source of up-to-date 'knowledge'. GPs' information-seeking could be exploited for guideline implementation. "*There is a lot of pressure to prescribe statins and I think we need to be sure we know what we are doing and how we are doing it*" (P8). Generally clinical guidelines were perceived as useful in improving the knowledge base and providing information to practitioners. "*Guidelines have helped me to improve my knowledge base from time to time*" (A4). This guideline-derived knowledge could then be used for changing behaviour, but also in many cases to 'confirm' the current good clinical practice: "*I must admit I only use them if I'm finding difficulty in the managing of particular patient. And would probably use them to just confirm*" (P10). Many GPs defined this approach towards guidelines as 'using' them, but not all had the same terminology: "*While I don't use [SIGN guidelines], I refer to them quite a lot. And they are extremely good*" (A3). The importance of understanding the meaning behind terminology was noted before when discussing 'evidence' (4.3.3). This represented another example.

VII.4. Ownership – local versus national guidelines

Interviewees expected the publication of national guidelines to overcome the problem of conflicting messages and "*too many separate groups producing guidelines*" (P10). They defined national guidelines as those produced either by a national authoritative body (e.g. NICE) or a body who had achieved nationally recognised status (e.g. BTS). "*For example the BTS [asthma] guidelines are national guidelines, everyone's doing the same thing, and I think people quite like that*" (P11).

Local ownership could increase the perceived credibility of a guideline and hence improve implementation. The interviewees argued that local ownership (through producing local versions of national guidelines or membership of guideline development groups) had an important facilitating effect on adherence. "*A meeting within a practice developing their own guidelines, perhaps based on national guidelines, and I think the key word is ownership. I think if doctors ... have ... some ownership of the decision making process, then they are much more likely to be successful*" (P2). "*An in-house protocol probably got the greatest credibility of all. I don't use external guidelines*" (A3). "*The biggest barrier, I think is not being locally owned*" (A8).

The process of local adaptation might also improve the chance of in-depth understanding of guideline recommendations as well as the underlying justifications for the recommendations. "*Osteoporosis guidelines we produced in [name of city], they certainly were useful, because we actually worked them up together as GPs*" (A8). The same might be achieved where GPs were members of guideline development groups. "*Because everybody participated in creating them, everybody is participating in the using of them. It is a key factor. It's not the government telling us what we should be doing*" (A6). "*I had contact with North of England guideline for non-steroidal anti-inflammatory drug for osteoarthritis. So [I] tend to use paracetamol or paracetamol-based analgesics first. ... Because I was a member of the working party for that guideline*" (P8). Obviously not all GPs could be members of guideline development groups and not even all of them could contribute to local adaptation. The sense of

ownership could be to some extent achieved if guidelines were produced so that GPs views were clearly represented in the process.

VII.5. Enforced implementation; medico-legal issues

Quality markers or indicators were considered as means of enforcing implementation of guidelines. Where used, they could be effective in behaviour change. "*[NICE Relenza guidance] I don't quite know how did they come to that decision but it's difficult not to pay attention to them because they are going to be national markers of what we were up to...*"(P5). Some perceived governmental targets (such as those in NSFs) as imposed. Those recommendations were seen as things that they 'had to' follow. "*We are trying to get National Service Framework for heart disease implemented. ... You've got to do it. If you don't you can't get your money or your quality markers. ... It's a stick, isn't it? That means you have to do that guideline, whether it's right or wrong*" (P7).

GPs thought some of these markers were not justified and were not based on evidence. "*If our performances are being measured by a certain percentage of our patients being on beta-blockers after MI, who's to say what percentage of patients should be on beta-blocker after MI? Because nobody on trials is of our patients. They've done trials on highly selective groups of people*" (A9). The negative feelings of GPs towards guidelines without their contribution were briefly encountered, while discussing previous sub-themes (4.3.4). GPs saw it insulting if people who had no responsibility for 'doing the job' produced guidelines for GPs. "*To me - to be honest – that's quite insulting because the people who actually are not doing the work and not involved with it, draw the guideline and I don't see how this expert body can expect it to be implemented*" (P5). Seeking GP support for the targets would ease their achievement. "*Try to persuade people that actually this is OK to follow. It is by us, for us, not dropped upon us*" (A6). Although 'imposed' guidelines achieved some results, especially in short term, they might have negative consequences (Lewin, 1951). 'Imposing' should not be the first option. Enforcement might reduce morale in general practice and should be limited to the vital areas and to the topics where other interventions proved ineffective.

Audit (and feedback) in general was perceived as a facilitating factor. Two reasons were expressed. First, GPs became aware of their performance as a result of audit and tried to improve it: "*our primary care development plan ... requires us to audit ... and it showed that we weren't doing very well with cholesterol. ... ... that was the biggest factor [in improving the service], the fact that we did an audit*" (P6). Second, as comparative practice data were distributed among their colleagues, they themselves tried to keep up. "*And that's a bit of peer competition I suppose. In a silly way, you don't want to fail in front of your peers ... it is a system that works: audit and then threats that those figures are going to be published or circulated*" (P6). Another study argued that this competition effects as a result of audit tended to have limited effects (Sheaff et al, 2003).

A further use of clinical guidelines could be in defining minimum standards of care and trying to achieve those minimum standards by clinical governance. One interviewee explained this approach in more details. "*They are not called guidelines anymore. They're called minimum standards. We've actually moved on from guidelines towards clinical governance. ... You are not trying to get local ownership of them; you say 'right, can we deliver this?' Rather than rewriting them for the practice, you actually say 'well, we can't find our hysterectomy patients, because we don't have a computer.' So you then [go] to PCG and say 'I need a computer'...*" (A8).

Medico-legal consequences of adhering to or ignoring a guideline could also act as enforcement. Some interviewees referred to this, mainly as an action that authorities might take if GPs did not follow guidelines. They did not sense strong pressure from potential medico-legal consequences instigated by health authorities. The other potential medico-legal liability was that as patients became aware of the guidelines, they could sue their GPs if they did not follow the guidelines. Again there was no widespread concern about it and no experience or observation was reported.

VII.6. Supporting implementation and implementation cost

Persistent and repeated distribution of guideline messages was an important implementation facilitator. Again they used BTS asthma guideline as example. "*... probably the number of times that I'm exposed to it. So for example the BTS guidelines, they hit you over and over again*" (A11). "*Repetition is another thing. If it's the first time that they've heard of the guideline, they are liable to ignore it. BTS was quite example of that. Because you got it from everybody you talked to about asthma. Reps talked about it*" (A5). Hearing the guideline message repeatedly could have a facilitating effect, if those different sources provide non-conflicting messages. "*That's also something very important. The same message is presented in different ways, but they are not conflicting messages*" (A6).

The idea of financial incentives and rewards for practices that improved the quality of their prescribing had some support. "*The rewards can be in terms of chronic management payments*" (A6). But some thought there were not enough financial incentives to follow guidelines. "*I need to be able to give you an incentive to change your habits from neproxen to ibuprofen and there aren't any. There are no real incentives. Often there is talk that financial incentives are useful, but there aren't really financial incentives available in order to do this*" (P2). PCOs' financial resources were perceived to be insufficient. Sometimes not all extra costs incurred from guideline implementation were met. An example was prescribing costs of statins if guidelines implemented: "*[PCO] get you thinking about statins guidelines, but they don't help you*" (A8).

GPs did not welcome cost-containing messages in clinical guidelines, especially if they could not clearly see the justifications. They needed to be convinced that guideline objectives were explicit and justified. "*I think clinical guidelines are very useful provided they are logical ... provided we can see the logic behind them. I think we tend to be suspicious that there is an alternative motive in producing guidelines which basically is to restrict amount of money spent on prescribing which may or may not be justified*" (P8).

There were times when the cheapest treatment option (with equal effectiveness) was not the best option from GP point of view. Minor side-effects, extra appointments required, number of times the drug should be taken by patients were among factors that contributed to shaping GP view towards guidelines recommendations. "*I think it's easy for a pharmacist to say prescribe four tablets a day, but it's easier for GP to say no you can just have one tablet a day which is minocin. And I think we are less worried about costs when we are dealing with teenagers with acne; four tablets versus one, there is no comparison*" (A4). This had a message for planning of cost-effectiveness studies. Far too often small but important costs to patients and GPs (e.g. time, convenience) were ignored in the studies. One interviewee provided another illustrative example of this issue: "*our practice guidelines [for depression] are about 3-4 years old now and I'm pretty sure aren't being followed at all. ... The guidelines in depression tended to say start with tricyclics, now most of the people we see with depression, are working and also driving a car. And ... tricyclics ... cause drowsiness*" (A9).

4.4. Discussion

4.4.1. Limitations and strengths

In recent years several qualitative studies of prescribing in British primary care have been published. The studies focused on different issues and problems such as change in prescribing (Armstrong et al, 1996; Allery et al, 1997), prescribing of new drugs (Prosser et al, 2003; Jacoby et al, 2003), patient-doctor relationship (Britten and Ukoumunne, 1997; Britten et al, 2000), variation in cost and prescribing (Carthy et al, 2000), implementing clinical guidelines (Freeman and Sweeney, 2001), implementing hypertension guidelines (Cranney et al, 2001), use of statins (Fairhurst and Huby, 1998), secondary prevention of CHD (Summerskill and Pope, 2002), effectiveness of computerised clinical guidelines systems (Rousseau et al, 2003), antibiotic prescribing (Kumar et al, 2003) and antithrombotic treatment for atrial fibrillation (Howitt and Armstrong, 1999). Some studies included groups outside primary care teams such as consultants (Allery et al, 1997) or patients (Britten et al, 2000). The qualitative study reported here was different from previous studies in two aspects. First, it had a unique sampling approach as it included GPs as well as academics of primary care. Second, the study asked about the views of the participants about implementing prescribing recommendations of clinical guidelines for five clinical topics; asthma, statins for prevention of CHD, menorrhagia, epilepsy and depression were directly covered in the interviews. The topics ensured variation in terms of complexity of the disease, availability of clinical guidelines, GP role in prescribing and the importance of prescribing in management of the problem (4.2.3).

The study had important limitations. The analysis, as well as data collection and transcription, was conducted by single researcher. This might have increased the subjectivity of the findings (Mays and Pope, 1995). To reduce this limitation, different sections of the results were presented in in-house and external seminars, advisory meetings and national and international

conferences (e.g. Rashidian and Russell, 2002a and 2002b). Based on the feedbacks received, the analyses were re-checked to ensure the findings were supported by the data and that alternative interpretations were also reported. The findings were also backed up with verbatim quotes from the interviews to help the readers examining the validity of the findings (Greenhalgh and Taylor, 1997). Quality of data was not audited by others. Errors might have happened while transcribing the data. Attempts were made to reduce the likelihood of errors. The tapes were listened to twice before transcribing and were transcribed with care. The validity of the findings could have been improved by feeding back the results of the analysis to the interviewees ('respondent validation', Mays and Pope, 2000; or 'member checking', Giacomini et al, 2000). This was not conducted because the first full-draft of the analysis suitable for feedback was prepared about one year after the interviews. It was perceived that the respondents were likely to have forgotten the content of the interviews and also might have modified their views. The delay in preparation of the full-draft happened because of the time required for the transcription and analysis of qualitative data, as well as the time that the researcher had to spend on preparation of the next stages of the project.

Some qualitative studies used random sampling (Dowswell et al, 2001; Harrison et al, 2003). Objective sampling approaches provide the chance of identifying trends and generalisable phenomena which may be missed in purposively sampled studies. Some random sampled studies were conducted alongside RCTs, comparing different trial arms (Harrison et al, 2003). That approach increased participation rate, which could otherwise be a threat to random-sampled qualitative studies.

The inclusion of two different groups of interviewees (GPs and academics of primary care) into the study improved the comprehensiveness of the study (Mays and Pope, 2000). It also enabled the researcher to look for systematic differences between the two groups. The differences in the views of the interviewees were highlighted when discussing different themes. Discordant views were more prominent when discussing particular examples (e.g. whether the BTS asthma guidelines were evidence-based or simple or whether the NICE Relenza guidance was useful). There was more congruence on the content and

general meaning of the themes. There were also extreme views, e.g. one academic of primary care (A8) thought that consultants should be kept out of any guidelines developed for primary care. The discordant views were not along the lines of the two different groups of interviewees (i.e. practitioners versus academics). The exception was on whether the guidelines resulted in change in behaviour (4.3.9). Generally there was more support among the non-academic GPs for the view that some guidelines caused changes in practice, while academics tended to be more doubtful of the cause of observed changes. Apart from the above example, discordant views reported under different themes were dispersed between the two groups. For example, in case of NICE Relenza guidance, one academic of primary care perceived it as something unrealistic (A9) and another one considered it as "something that GPs can stick to" (A2).

Obtaining views of other important groups such as practice nurses, consultants and practice managers might have increased the comprehensiveness of the study, but these were beyond the stated objectives. The sampling approach and the inclusion of academic practitioners might have caused biases (selection bias) in the results, as the main criteria were to invite those with active experience of guideline implementation. Therefore positive views towards guidelines might have been over represented. Despite these concerns, many negative comments and views towards clinical guidelines and their use were received from the interviewees, which were perhaps the outcome of their experiences. Focus on specific clinical conditions might have reduced the generalisability of the findings to other clinical conditions. This was, however, intended and provided the chance to obtain more detailed views about the target issues.

The framework method helped structuring the process of qualitative analysis and hence making it more objective. It also enabled explicit incorporation of a priori theoretical and empirical reasoning into the analysis (Pope et al, 2000). The purposive method of sampling and focus on guidelines for specific clinical conditions helped this study to go further in exploring GPs' views. The study identified and reiterated many factors that hindered or facilitated implementation of guidelines to change prescribing in primary care. More importantly it mapped and defined the inter-relationships between these

factors. It reinforced findings of some previous studies, challenged others and raised new questions. The findings were confined to the merits and limitations of qualitative studies (Murphy et al, 1998). The study strongly supported the theoretical standing that '*despite their specificity, clarity, and credibility, guidelines are unlikely to produce significant change in clinical practices without carefully designed and executed programmes to achieve their implementation*' (Mittman et al, 1992, p 421).

4.4.2. Evidence, usefulness and relevance

Subsection 4.3.3 reported the participants believed guidelines should be based on evidence as an essential feature. The thematic framework suggested that evidence base, flexibility and change formed the credibility of content of a guideline, acknowledging the limitation of research evidence in convincing GPs to implement a guideline. In a well conducted descriptive study, Grol et al identified a few attributes of guideline recommendations that influenced implementation (Grol et al, 1998). Among those were (positively worded) recommendations being 'non controversial and compatible with current practice', 'precise and specific' and 'not demanding change of routines'. Being evidence-based was another important attribute, but its weight was equal to 'not having consequences on management'. This analysis supported Grol et al's finding in two directions. First, the evidence base of guidelines had limited role in enhancing implementation. Second, perceived relationship between guideline and change in practice was important in implementation.

In this study, participants had different views on what constituted the best evidence to base their practice on. This raised the question of whether it was possible to propose a new level of evidence for general practice, while it still remained methodologically credible. Green (2000) encountered the same phenomenon while studying road safety issues through interviews with a mixed sample, which included managers, health promotion staff, road safety officers and police officers. She saw although they all agreed that decisions should be based on evidence, they did not agree on the nature of evidence (Green, 2000).

She offered two themes to explain the differences: that evidence was 'situated within specific agendas' and that it was 'constructed through professional practice'. This qualitative study supported the first theme. GPs found evidence less useful if its agenda was different from their agenda. The second theme, construction of evidence through professional practice, did not explain the variations observed in the data. Within the sample of interviewees it was unlikely to relate differences in views to different professional practices or their access to different sources of information (e.g. publications).

The presence of variation in understanding of 'evidence' raised some questions. Was evidence-based medicine becoming a term that people used for different purposes and different meanings ('agendas')? To what extent research evidence was important for implementation of guidelines' prescribing recommendations (i.e. could it be quantified)? Wood et al (1998) concluded that 'evidence was important, but not that important', adding, '*one requires precision in talking about evidence*' (p1735). This statement was very similar to what found in this study. The term 'evidence' could play the role of a jargon and could support opposing positions. Also it *might* or *might not* apply to specific phenomena depending on one's choice and viewpoint: a small trial might be seen as credible evidence to some GPs and as irrelevant and non-helpful piece of research to others. Similarly others argued that successful implementation was not based on the rigor of the evidence, but on three variables: the evidence, the context, and facilitation of implementation process (Kitson et al, 1998). Wood et al (1998) went even further and argued:

> *there are simply (re)constructions of evidence able to support almost any position. ... The nature of evidence is ambivalent. It is constructed into debates and controversies, which are often equally supportive of opposing viewpoints*" (pp 1735 and 1737).

Although Wood et al might be able to provide 'evidence' (i.e. examples) for the above claim, it did not seem to be accurate. There were many examples where evidence overwhelmingly supported a specific clinical practice over the competing alternatives; and some were presented in the findings.

Other interesting terms were 'relevance' and 'usefulness'. Some GPs found the 'evidence' as irrelevant to their individual patient condition or simply useless. This was also noted in previous studies: '*the quality of evidence upon which guidelines are based is frequently criticised, on the grounds that most trials are hospital based ... and always exclude a proportion of patients*' (Dowswell et al, 2001, p 121; see also Schwartz et al, 1989 and Cranney et al, 2001). Therefore,

> *unless academics can learn to understand the clinician's definition of evidence and effectiveness, they cannot hope to have a major impact on the latter group's prescribing practices* (Schwartz et al, 1989, p 581).

GPs uncertainty in what was the best evidence to base a recommendation on resembled the difficulties the methodologists face in grading clinical guideline recommendations. The traditional approach was to assign the highest grade to the recommendation if they were based on evidence from good quality meta-analyses or RCTs. This traditional wisdom was challenged and some argued that grading of recommendation was not just about the evidence base of recommendation (GRADE Working Group, 2004). The new approach suggested consideration of a few other factors as well as quality of evidence: 'balance of benefits and costs', 'balance of benefits and harms', 'strength of recommendation' and its 'relative importance'. The qualitative study supported inclusion of other factors when grading evidence. The problem would now be that recommendations might become more subjective and potentially prone to new criticisms.

Interviewees referred to flexibility as a factor that could accommodate both research evidence and real practice together. Some researchers adopted the same position (Graham et al, 2000b). They challenged the assumption that variations in care processes and outcomes reflected poor practice in primary care. Studying the validity of heart failure guideline for primary care, they offered their 'conundrum' as: '*clinical guidelines must be both clinically applicable to individual patients and whole populations'* (p 953). However, this might not be an easy task to achieve and probably 'a bridge too far'. Again the definition itself required more clarification and boundaries were not precise. In order to cover that level of comprehensiveness in guidelines, exhaustive and

detailed documents would be required. The resulting guideline would perhaps be more complicated than what many GPs expected to see, as discussed in details under 'simplicity' (4.3.5).

Change in guidelines featured as an important factor that could result in reducing the credibility of content of a clinical guideline. Change in guidelines (for example through updating) negatively affected implementation through impeding clinicians' progress through the learning curves (Maisonneuve and Tiiu, 1999), although this was probably more relevant to complicated procedures than to prescribing.

4.4.3. Guidelines and practice improvement

Answering the question of whether guidelines changed prescribing was not the objective of the qualitative study. This was better done through systematic reviews of trial evidence (e.g. look at Grimshaw and Russell, 1993; Grimshaw et al, 1995; Grimshaw et al, 2004; also see Chapter 2) or qualitative evidence (e.g. see Ferlie et al, 2001). In particular, guidelines alone had limited impact on behaviour change. Allery et al (1997) studied why GPs changed their clinical practices (including prescribing) and found a small role for guidelines. However, closer look at the way Allery et al analysed their data raised some questions over the approach they used for categorisation of reasons for change. It seemed that guidelines might have contributed through other factors that Allery et al identified. GPs used guidelines for a variety of purposes including for improving communication with patients. Similar to this analysis, others found that '*the GPs appeared to be more interested in using evidence to reassure patients and answer any queries they may have, than change their practice based on it*' (Mayer and Piterman, 1999, p 631).

Armstrong et al (1996) found that implementation was dependent on the experience of GPs using drugs on their patients. They identified three models of change in prescribing: accumulation model, challenge model and continuity model. The findings reported in this chapter were in line with accumulation

model (e.g. statins) and continuity model (e.g. asthma). Armstrong et al did not observe much change in GPs prescribing, regardless of many potential sources of encouragement for changing prescribing.

Several studies and reports tried to highlight attributes of good clinical guideline (Field and Lohr, 1990; Cluzeau et al, 1999a; Cluzeau et al, 1999b; Grilli et al, 2000; Graham et al, 2000a; The AGREE collaboration, 2001; The AGREE collaboration, 2003). Earlier papers did not include the issues relevant to application and implementation of guidelines among the attributes. This was rectified when an international tool for guideline appraisal was developed (The AGREE collaboration, 2003). One domain of the appraisal tool was called 'application' and covered three items of 'cost implications', 'organisational implications' and 'audit criteria' (The AGREE collaboration, 2001). The major limitation of the appraisal tool was that it was more concerned with the process of guideline development than it was with the content (or quality) of clinical guidelines. However, an empirical study suggested that the tool related with some aspects of credible content of clinical guidelines (Irani et al, 2003).

The Institute of Medicine (Field and Lohr, 1990) identified eight attributes of 'good practice guidelines'. Those attributes were validity, reliability and reproducibility, clinical applicability, clinical flexibility, clarity, multidisciplinary process, scheduled review and documentation (ibid, p 59). Almost all of these attributes were highlighted by the participants in the qualitative study, with the exception of scheduled review. Reliability and reproducibility were also indirectly mentioned. The study, however, implied that not all attributes were given the same weight in GP view. Some GPs gave further weight to attributes that others saw as less essential. The findings also suggested the addition of one further attribute to the list. Guidelines should be clear in their stand against current practice and also the potential changes they propose. The new attribute might help clinicians to obtain better understanding of why the particular guideline was considered as required, which was far too often unclear to clinicians. It would also help clinicians and managers to plan resources, the time required for implementation and the actions to be undertaken in response to the guideline. We called this attribute 'planning for change'.

The other, rather important finding of this qualitative study was that even if all attributes of 'good clinical guideline' were present, guidelines' success in improving quality of care required more conditions to be met. The findings suggested that the attributes of good clinical guideline might be more attractive to methodologists and developers of clinical guidelines than they were to clinicians. In reality clinicians were not equipped with reliable tools and methodological knowledge to test the 'validity' or 'reliability and reproducibility' of clinical guidelines. Instead, as the findings of the qualitative study suggested, clinicians tended to use credibility of source of guideline and 'respected others' recommendations as proxy for those attributes. The AGREE instrument could overcome the problem, at least in part (The AGREE collaboration, 2003), and therefore it might be useful asset for clinicians and other members of primary care teams. The AGREE instrument was used in two recently conducted workshops in London (Rashidian, 2004a) and Cambridge (Rashidian, 2004c) in which the participants appraised clinical guidelines that they had different levels of familiarity with. Although none of the participants had any previous experience with the instrument, they were able to perform the task. It is yet to establish the validity of judgements made using AGREE, but as mentioned previously at least one study suggested that guidelines which obtained higher score with AGREE seemed to be of higher content validity (Irani et al, 2003).

4.4.4. Credibility of source

It has been argued that credibility of source of a guideline, or the messenger who is transferring it, are important elements in guidelines' adoption by practitioners (Raisch, 1990a; Rogers, 1995b). The qualitative study provided further evidence for the importance of credibility of source. In terms of messengers for transferring of guideline, Rogers thought that physicians considered 'near-peers' as more trustworthy than others (Rogers, 1995b). The analysis suggested that GPs generally considered other GPs as trustworthy, but this was not always the case for hospital consultants. Like other studies consultants were viewed as valuable sources of advice and support (Carthy et al, 2000), but it depended on context, and negative views were expressed at the

same time. Others studies suggested that credibility of evidence or source of evidence was related to professionalism (Bradley, 1992). This could explain somehow the negative views of GPs towards guidelines that were produced or promoted by consultants for GP use. Development of clinical guidelines for GPs by hospital consultants is still common practice in the UK (Bowens et al, 2001). Sometimes consultants spend as little as one hour on developing guideline for GPs (Bowens et al, 2001). The interviewees strongly believed that GPs should be directly involved in the development or local adaptation of clinical guidelines for primary care.

Rogers (1995) viewed diffusion of innovation as a social process, so that physicians sought '*information from peers, usually those who have previously adopted the innovation of interest'* (p 326). With this in mind, it was easier to understand why consultants were not always 'near-peers'. Consultants did not work in the same setting as GPs and therefore did not adopt the innovation in the way that GPs did. As Greer (1995) put it, appreciation of contextuality was wisdom: '*the many factors that impinge not on theory but on practice*' (p 328).

The study suggested GPs were frustrated not only because of the number of guidelines they received (Hibble et al, 1998), but also because of the perceived lack of relevance and quality of many of those guidelines. The participants felt that some secondary based guidelines did not relate well with their primary care practice. Evidence suggested that many guidelines suffered from poor quality regardless of who produced them (Shaneyfelt et al, 1999). A careful review of 431 guidelines from specialty societies suggested that very few met all criteria of good guidelines (Grilli et al, 2000).

The BNF had been shown as a credible source of evidence (Salisbury et al, 1998; Carthy et al, 2000). The study supported the inclusion of clinical guidelines in the Formulary to improve their credibility and implementation. Apart from BNF, there was limited support that for publications in other sources as an implementation tool. This reflected the findings of two previous studies (Timpka et al, 1989; Armstrong et al, 1996).

4.4.5. Patient pressure

When studying the reasons for 'non scientific prescribing', Schwartz and colleagues (1989) found out that 46% of physicians expressed 'patient demand' as the underlying factor. A further 24% mentioned that they used non-effective drugs because of 'placebo effect', implying it was the preferred drug for the patient (Schwartz et al, 1989). The qualitative study did not imply that patient preference were the main factor of non evidence-based prescribing (4.3.6). There could be two reasons for this difference. First, the qualitative study was different in approach and main questions from Schwartz et al study. They had identified physicians that prescribed certain drugs more than recommended and then had asked them why they did so. Second and probably more important difference was because of setting. Physicians participating in Schwartz et al (1989) expressed that they considered patient expectations because of the predominantly fee-for-service nature of health services in the US (at the time), which was different from the NHS: '*many feared that ... a failure to accommodate patients in prescribing could mean loss of business and reputation*' (Schwartz et al, 1989, p 579). The fee-for-service market might also explain why 1975 study in Iran concluded that inappropriate antibiotic prescribing was due to patient pressure (Amidi et al, 1975). Nonetheless, the finding that patient expectation was not the main underlying factor of non-evidence-based prescribing was in line with what being reported in previous reviews (Horder et al, 1986).

Other studies in Australia (Cockburn and Pit, 1997) and England (Britten and Ukoumunne, 1997; Britten et al, 2000) identified that patients were more likely to receive prescription if GPs perceived that the patients expected it. These studies were performed in non fee-for-service health systems, and still GPs were likely to prescribe according to perceived patient pressure. A recently published study of the importance of patient pressure in British primary care had interesting findings (Little et al, 2004). They measured patients' expectations before the visits and doctors' perceptions at the end of the visits. They found that the perceived patient pressure was independent predictor of prescribing, even when perceived medical need was controlled for (Little et al, 2004). The impact of (real) patient expectation was less prominent and inconclusive. Little

et al study was supportive of the TPB notion of subjective norm (Ajzen, 1991). The theory asserted that social pressure affected intention and subsequently behaviour if it was 'perceived' by the individual, whether or not it reflected the actual social pressure.

A UK study of factors associated with costs and variation in prescribing interviewed seventeen GPs in Avon and reported excessive and unrealistic requests for prescribing from patients as underlying reasons for low quality prescribing (Carthy et al, 2000). There were differences between Carthy et al's and the qualitative study presented here. First, Carthy et al assessed antibiotic prescribing, where GPs are urged to prescribe less. The present study focused on prescribing for predefined clinical areas in which guidelines were generally urging for more proactive prescribing of specific products. Second, in the present study, widely-known clinical guidelines for asthma treatment and CHD prevention could have helped GPs in better communication with patients and in convincing them. Also in light what reported by Little et al (2004) it was likely that the perceived pressure for low quality prescribing (Carthy et al) might not have reflected patients' real expectations.

4.4.6. Local initiatives and national priorities

The interviewees generally supported in-house and local adaptation of national guidelines. Local initiatives, especially at practice level, might explain some of variation in quality of care received by patients (Griffiths et al, 2001) and were more likely to succeed in implementation (Fairhurst and Huby, 1998). Social influence theory advocated use of participatory guideline development to increase ownership and implementation (Mittman et al, 1992). The issue of locally developed guidelines was discussed in Chapter 2 (2.3.6 and Table 2.1) and in 4.3.9. A systematic review of clinical guideline implementation programmes in primary care reported a success rate of four out of nine for local guidelines, and one out four for national clinical guidelines (Worrall et al, 1997). Despite that a review of rigorous interventions for guideline implementation identified four studies, of which only two studies found participatory guidelines

effective (Grimshaw et al, 1995). The overview of systematic review concluded that participatory guideline development was unlikely to be cost-effective (2.3.6). Local adaptation, however, could be used for setting standards of care and local targets (4.3.9) and might facilitate 'translation of evidence':

> *It follows, therefore, that research is rarely, if ever, self-evident to the practitioner, but varies according to the context within which it is received. ... In promoting an innovation or piece of research evidence we are not dealing with the uncomplicated dissemination of findings to a largely passive and receptive audience - a simple problem of ``putting theory into practice" in the hackneyed sense of the phrase - but with the question of reconnecting research with its supplementary other: practice. The key point here is how evidence is translated within the assumptive world of practitioners* (Wood et al, 1998, p 1734).

Despite benefits of local initiatives, national priorities should be taken into account. Making priorities based on local consideration might result in unnecessary focus on areas that were already in satisfactory status (Haines and Jones, 1994). In recent years in England, national priorities often emerged through NSFs and more specifically via guidance and guidelines issued by the NICE. Credible and evidence-based performance indicators might improve implementation of NSF recommendations and NICE guidelines and give authority to monitor change (Wilkinson et al, 2000). However, it might be the case that most changes happen in data recording. Also the quality gap might increase as practices with better infra structure and quality of service would respond more rapidly to the initiatives (Wilkinson et al, 2000) and get rewarded. The analysis of the interviews implied that enforcing guideline implementation could be successful in limited instances, could be short lived, and might hinder long term success of implementation programmes. According to the field theory (Lewin, 1951) increasing deriving forces without addressing resisting forces leads to increased tension without much achievement. Approaches that includ imposing guidelines might '*force practitioners to develop avoidance strategies, but not change actual prescribing intentions*' (Raisch, 1990b, p 540).

Although financial, managerial and other pressures might be necessary to change behaviour in special circumstances, they should not be used as a routine. The study and previous research indicated that practitioners interpreted these pressures or objectives with suspicion (Mayer and Piterman, 1999). Also

the findings supported previous studies that physicians did not welcome cost-containment motives for change in prescribing (Raisch, 1990b). The respondents to a survey of 1549 physicians in the USA found it unacceptable and ethically incorrect if payers enforced guidelines. This was especially true when the guidelines included cost containment motives (Sulmasy et al, 2000). Cost containment motives should be explicitly defined in the guidelines (Mayer and Piterman, 1999), especially as it was not uncommon for guidelines to aim for reducing costs (O'Brien et al, 2000).

4.4.7. Practice nurses, prescribing advisers and reps

Some interviewees considered practice nurses as more active in guideline implementation than GPs. Other studies of practice nurses in the UK had found that nurses were generally welcoming towards clinical guidelines (Harrison et al, 2002; Puffer and Rashidian, 2004) and encouraged GPs to be likewise (Cranney et al, 2001). The participants also found nurses positively influential on their implementation of clinical guidelines. Also the participants had positive views on the effects of nurse-run mini-clinics on guideline implementation. Four-year follow-up of an RCT demonstrated that CHD prevention clinics were linked to improved quality of patient care, potential life saving and fewer cardiac arrests (Murchie et al, 2003).

The study suggested that prescribing advisers had only limited roles in implementation of guidelines, similar to another British study (Salisbury et al, 1998). This was despite the theoretical support (Mittman et al, 1992) and empirical evidence for positive effects of educational outreach visits on prescribing (see 2.3.8). However more recent studies found the effectiveness of educational outreach variable (2.4.1). Two recent British trials indicated that the effects of prescribing advisers were likely to be minimal and limited to certain conditions (Watson et al, 2001; Freemantle et al, 2002). These findings and the interviewees responses might be explained by social influence theory, which implied that educational outreach visits were likely to be more successful when advisers were known and respected by target GPs (Mittman et al, 1992).

Interviewees acknowledged the role of drug company reps in the implementation of BTS asthma guidelines and less prominently in the diffusion of CHD risk estimation charts. It was suggested that this sort of influence tended to be indirect (Rogers, 1995b). A British RCT failed to demonstrate any effects from drug company reps' visits (Freemantle et al, 2000). Similar to the views expressed by the interviewees, previous research demonstrated that many members of guideline development groups had relationship with drug industry (Choudhry et al, 2002). The participants found this a threat to the validity of clinical guidelines.

4.4.8. Information technology; medico-legal issues

Previous research concluded that computerised systems may not be that effective in improving implementation of guidelines for a chronic disease like diabetes (Hetlevik et al, 2000), mental health care (Lewis et al, 1996) or even lipid lowering prescribing (Hobbs et al, 1996). The benefits of computerised systems tend to be limited to certain areas of care including screening and prevention (see 2.3.10). At the time that this study was conducted there were limited research on the effectiveness of computerised clinical guidelines (e.g. PRODIGY) in improving quality of care in British primary care. A recent cluster randomised trial and a follow up qualitative study concluded that the computerised system was unsuccessful in improving implementation of clinical guidelines for asthma and angina (Eccles et al, 2002; Rousseau et al, 2003). These studies supported the finding that incorporation of clinical guidelines for complicated clinical conditions (e.g. hypertension, asthma) into software packages would have a limited impact on implementation.

Respondents did not perceive that guidelines were causing them more susceptible to medico-legal claims. This was in line with what predicted in a previous systematic review of evidence (NHS Centre for Reviews and Dissemination, 1994). They referred to the Bolam test as a basis for identification of clinical negligence cases in court (Jones, 2000). The Bolam test

used the criteria of common professional practice rather than whether the practice was based on evidence. Based on this test professional non-adherence to a guideline was negligent only if it was shown that the guideline was so widely used and accepted by the professional community that no reasonable skilled professional failed to comply with the recommendations (NHS Centre for Reviews and Dissemination, 1994). The Bolam test has been challenged in more recent court rulings (Jones, 2000; Samanta and Samanta, 2003). Despite this the interviewees did not perceive medico-legal issues a major source of discontent for guidelines.

4.4.9. Thematic framework, TPB and the taxonomy of interventions

The qualitative study findings were used to update the taxonomy of interventions presented in Tables 2.1 and 3.3. A further column was added to the original table to present whether GPs perceived the intervention as 'effective', with 'limited effect' or 'not effective'; or whether there was no mention of the intervention in the interviews ('not stated'). One intervention (essential drugs programmes) was specific to developing countries context and therefore was noted as 'not applicable' (Table 4.2). 'Not stated' category might imply that the intervention had limited application for changing prescribing behaviour, it was not effective or it was not widely experienced in the UK. It was assumed that interventions with no evidence of effect and 'not stated' in the interviews were unlikely to be important assets for behaviour change. There were important agreements between the findings of the systematic review and the qualitative study. This might reflect the sampling of the qualitative study and that the respondent might have been informed of the literature. There were also important disagreements. For some intervention with no evidence of effect, the qualitative study provided support. This could be informative since not all interventions could be assessed in experimental or quasi-experimental studies. The comparisons were summarised in Table 4.3.

TABLE 4.2. SAGE TAXONOMY OF INTERVENTIONS, VERSION II. STRATEGIES FOR IMPROVING PRIMARY CARE PRESCRIBING: EFFECTIVENESS EVIDENCE PLUS THE QUALITATIVE STUDY FINDINGS

	Evidence from the overview of systematic reviews					Opinions of the interviewees
	Intervention	**Relevant theories**	**Effect**	**Cost**	**Durability**	
Competence oriented	CME (interactive)	Adult learning; TPB	++	Medium	?	Not stated
	IP education	Adult learning	?	Medium	?	Not stated
	Mailed printed material	Social influence; TPB	+/0	Low	?Short	Not effective
	Mailed national warnings [$]	Social influence	+	Low	Medium to long	Not stated
	Participatory guideline devel.	Social influence; TPB; health education	+	High	?Medium to long	Effective
	IP shared care (consult-liaison)	Health education; TPB	+/0	?High	?	Effective / context specific
	Mass media	Diffusion; TPB; social influence	+	Low	?Medium	Not stated
Performance oriented	Audit and feedback[$]	TPB; social influence; stages of change	+	Low to medium	?	Effective
	Reminder systems	TPB	?+	Low to medium	Short	Limited effect
	Educational outreach	Diffusion; social influence; stages of change; TPB	?++	Low to high	Short to long	Limited effect / context specific
Social influence	Peer review	Social influence; TPB; Diffusion	?	?Medium to high	?	Effective
	Patient mediated	Social influence; TPB	+	Low to medium	?	Effective
	Local opinion leaders	Diffusion; Leadership; TPB	?	?Medium to high	?	Stated once / effective
	CQI [$$]	Management theories	?	?Medium to high	?	Not stated
Physical support	Practice support	Management theories; TPB	?	?	?	Effective
	Essential drugs programmes [%]	Management theories	?+	?	?	N/A
Financial incentives	Financial incentives	Economic; social influence	?+/-	Low to high	?Short to medium	Limited effect
Non-voluntary	Reimbursement and budgetary policies	Economic theories; TPB	?+/-	?	?	Limited effect
	Rules, obligations	Economic and management; social ecology	?	?	?	Effective
	Restricted formulary	Economic; social ecology	+	?	?Medium to long	Effective
	De-registration / reclassification	Economic; social ecology	?+/-	?	?	Not stated

IP: Inter-professional, ++: strong evidence suggests positive (intended) effects, +: limited evidence suggests positive (intended) effects, +/0: variable effectiveness, ?: no evidence of effect, ?+: no evidence of effect, however less reliable evidence suggests positive effects, ?+/-: no evidence of effect/ likelihood of positive and negative (intended and unintended) effects, $ it usually also has an element of social influence, % it usually incorporates competence oriented approaches, $$ it usually has elements of physical support and competence oriented approaches.

TABLE 4.3. COMPARISON OF THE FINDINGS OF THE OVERVIEW OF SYSTEMATIC REVIEWS AND THE QUALITATIVE STUDY ABOUT THE EFFECTIVENESS OF DIFFERENT INTERVENTIONS

Intervention	Evidence of effect	Perceived effectiveness (extracted GP opinions)
Congruent findings		
Participatory guideline development	+	Effective
Audit and feedback	+	Effective
Patient mediated	+	Effective
Restricted formularies	+	Effective
Reminder systems	?+	Limited effect
IP shared care	+/0	Effective / context specific
Incompatible findings		
Educational outreach	?++	Limited effect
Mailed printed material	+/0	Not effective
Interventions with no evidence of effect that were perceived effective		
Peer review	?	Effective
Local opinion leaders	?	Effective
Practice support	?	Effective
Rules, obligations	?	Effective
Interventions with no evidence of effect and not stated in the interviews		
IP education	?	Not stated
CQI	?	Not stated

++: strong evidence suggests positive (intended) effects, +: limited evidence suggests positive (intended) effects, +/0: variable effectiveness, ?: no evidence of effect, ?+: no evidence of effect, however less reliable evidence suggests positive effects, ?+/-: no evidence of effect/ likelihood of positive and negative (intended and unintended) effects

Identification of the qualitative themes was an evolving process and matured during the process of the analysis, in line with the framework methodology (Ritchie and Spencer, 1994). The thematic framework started with nine main themes and reduced to seven (Table 4.1). For example analysis revealed that all of the sub-themes of 'practitioner related attitudes' were somehow relevant to the other main themes, hence this theme was dropped from the final framework. The themes also were not completely exclusive and there were degrees of overlaps. For example, there were difficulties in categorising the items relevant to patient-doctor interaction under the main themes. In the end, it was decided to split the items between two different themes of 'credibility of content of clinical guideline' under 'flexibility' and 'influential people in implementation' under 'patients'. A similar situation happened when dealing with items explaining the role of secondary care professionals in implementation of guidelines in primary care. In this case, the items were split between 'credibility of source of clinical guideline' and 'influential people in implementation'. The other example was 'publishing clinical guidelines in respected source' as a sub-theme. Although it could have been discussed under 'dissemination strategy', it was decided to put it under

'credibility of source of clinical guideline'. It was argued that guidelines were usually published by the developers, while other dissemination activities were normally performed without the developers' involvement or their knowledge. On the other hand, the sub-theme of 'ownership' was categorised under 'dissemination strategies'. Although local ownership of guideline improved credibility, it was usually undertaken as part of the dissemination process.

TABLE 4.4. SIMPLIFIED THEMATIC FRAMEWORK. FACTORS THAT MAY EXPLAIN VARIATION IN EFFECTIVENESS OF INTERVENTIONS TO IMPROVE PRIMARY CARE PRESCRIBING

Theme I: Credibility of content of clinical guideline	
Evidence-based	Flexible
Reflect change in practice / Avoid change in recommendations	
Theme II. Credibility of source of clinical guideline	
National professional bodies e.g. BTS (reputable)	Secondary care (not reputable)
National governmental bodies e.g. NICE (reputable)	Pharmaceutical companies (not reputable)
Published guidelines in respected sources (e.g. BNF, Clinical Evidence, BMJ, Lancet)	
Theme III. Presentation of clinical guidelines	
Simple	Systematic presentation
Theme IV. Influential people in implementation	
Patients (barriers and facilitators)	Practice nurses and primary care team (usually facilitators)
Consultants (barriers and facilitators)	PCOs, pharmacists and prescribing advisers (limited effects)
GP colleagues (barriers and facilitators)	Drug companies and reps (barriers and facilitators)
Theme V: Organisational factors	
Practice characteristics (mini-clinics)	Time, workload and information overload
Information technology (effective in limited circumstances)	Cost and expenditure
Availability of required resources	
Theme VI. Disease characteristics (where guidelines for GPs could be less effective)	
Treatment is secondary care based	Rare or 'simple' disease
Difficulty of diagnosis	
Theme VII. Dissemination strategy	
Planning Implementation	Ownership – local versus national guideline
Perceived need, experience and knowledge	Enforced Implementation
Supporting implementation and implementation cost	Medico-legal issues

Simplified version of the thematic framework was presented in Table 4.4. The sub-themes were re-worded to ease the understanding. Short explanations (in brackets) were added to the sub-themes to demonstrate how the sub-themes might influence implementation. The simplified thematic framework was incorporated in a secondary analysis of the data focusing on the relationships between different themes. The results of this secondary analysis were used to develop a simple model for implementing prescribing recommendations within guidelines in primary care. The model is presented in Chapter 5.

The thematic framework was developed following deductive (extracting sub-themes and items from theories and literature) and inductive (using data to develop the framework) approaches. Theory of planned behaviour was the main theory used for the development of the thematic framework. At superficial level, the final framework (Tables 4.1 and 4.2) might have little resemblance to the TPB and its constructs (3.2.6 and Fig 3.3). However the framework substantially benefited from the TPB. Among the main themes, Theme IV (influential people) was derived directly from the TPB's notion of 'subjective norm'; and the corresponding sub-themes reflected important groups that formed the salient normative beliefs of the interviewees. Other themes were also closely correlated with TPB constructs. Themes I and II (credibility of content and source) affected GPs' attitudes towards implementing guidelines. Themes III (presentation), V (organisational factors) and VII (implementation strategies) were more closely related to the TPB concept of 'perceived behavioural control'. Sub-themes identified as 'organisational factors' directly influenced guideline implementation, resembling the direct link between perceived controls and behaviour in TPB. The correlations between the themes and TPB construct were not exclusive. For example, 'credibility of source' could also affect 'subjective norms'. Only theme VI ('disease characteristics') did not seem to correlate with TPB constructs. 'Disease characteristics' was developed mainly as a 'screening' theme (5.3.1). The sub-themes could be used for identifying circumstance for which clinical guidelines for GPs might not be very effective.

The results of the qualitative study were used for the next stages of the project. Among five clinical conditions directly assessed in the study, only two conditions (asthma and statins for CHD prevention) had widely known clinical

guidelines and hence were suitable for use in the surveys. The qualitative study was also used for the identification of salient beliefs of GPs about implementing prescribing recommendations of clinical guidelines for asthma and CHD prevention. These beliefs were incorporated into the TPB questionnaires. Theory of planned behaviour was then directly assessed. The results are presented in Chapter 7.

McFarlane's law: '*when conflicting theories co-exist, any point on which they all agree is the one most likely to be wrong*'.

McFarlane, 1984, p 253

Chapter 5. SAGE model for implementing clinical guidelines in Primary Care Organisations in the NHS

5.1. Introduction

5.1.1. Background

Many (Grimshaw and Russell, 1993) have advocated the use of clinical guidelines to diffuse evidence-based practice. However more recent reviews (e.g. Oxman et al 1995; Davis and Taylor-Vaisey 1997; Bero et al 1998; NHS Centre for Reviews and Dissemination 1999) have shown that changing professional behaviour is difficult. Nevertheless some interventions are more successful than others (Wensing et al, 1998).

Against this background clinical governance was introduced (Department of Health, 1998). Since then the role of PCOs has been expanded to include quality improvement (Department of Health, 2001c). In this role they can exploit national targets (e.g. from National Service Frameworks – NSFs), guidelines and guidance (e.g. from the National Institute for Clinical Excellence – NICE) and quality reports (e.g. from the Commission for Health Improvement).

Nevertheless it is difficult for PCOs to improve quality of care through proactive implementation of guidelines. Given their limited resources and the

increasing cost of prescribing, extra costs are not welcome. However PCOs have some advantages. First they have more power to influence prescribing initiated in secondary care. Secondly, because PCOs scrutinise prescribing costs, GPs may be more willing to adopt evidence that reduces costs.

5.1.2. Existing models for implementing guidelines & changing prescribing behaviour

Understanding and synthesising the evidence on modifying physicians' behaviour is difficult. Stocking (1985) was one of the first to study the diffusion of innovation in British health care. Greer (1988) and Phelps (2000) saw the problem as one of information diffusion and others as one of research utilisation (Logan and Graham, 1998). Others focused on specific physician behaviours like prescribing (Raisch *1990a & 1990b*) and preventive care (Walsh and McPhee, 1992) or on specific underlying factors (Robertson et al, 1996). Implementing evidence-based practice and clinical guidelines is the focus of other publications (Rogers, 1995b; Grol, 1997; Kitson et al, 1998; Thorsen and Makela, 1999). These and others have proposed frameworks for implementing evidence-based practice (Stocking, 1985; Greer, 1988; Logan and Graham, 1998; Baker et al, 1999; Ashford et al, 1999; Phelps, 2000).

Previously systematic reviews have been used to generate an educational taxonomy of guideline implementation (Lomas and Haynes, 1988; Grimshaw and Russell, 1994; Wensing and Grol, 1994; Le Grand et al, 1999). Similarly the overview of systematic reviews was exploited to update taxonomy of interventions for prescribing behaviour in primary care (Table 2.1). Here the results of the secondary analysis of the qualitative data are used to generate a model of guideline implementation focusing on a single behaviour (prescribing) in a specified setting (primary care).

5.1.3. Theories of behaviour change

Grol (1997) recognised the importance of studying psychological theories of guideline implementation. Among theories for explaining behaviour in health professionals, social cognition approaches were focused on because these have more potential to explain guideline implementation (Bandura, 1986; Ajzen, 1991; Conner and Sparks, 1996). Among those, the TPB asserts that 'individuals make behavioural decisions based upon consideration of available information' (Conner and Sparks, 1996). Consistent with this Mittman et al (1992) described social influences on guideline implementation. The TPB argues that 'subjective norms' are among main predictors of intentions, and 'perceived behavioural controls' mediate other external influences on intentions and behaviour. Other factors that were considered in developing the model included credibility of evidence source (see 4.3.4.; also Filed and Lohr, 1990), marketing strategies (Rogers, 1995b) and organisational factors (Berwick, 1996). Coercion may also explain variations in practice (Grol, 1997).

5.2. Methods

This is a model of guideline implementation for PCOs across Britain. Though it is based on theories of behaviour change, technical terms typical of the behavioural sciences are excluded from it. Rather than make it exhaustive, it also excludes esoteric elements with less practical relevance.

The qualitative study comprised 25 semi-structured, in-depth interviews with GPs and primary care academics (Chapter 4). They were conducted in late 2000 and early 2001. Data was analysed using the framework method (Ritchie and Spencer, 1994; Ritchie et al, 2003). In principle the model was based on these interviews. To make the resulting model of guideline implementation for

prescribing in primary care more comprehensive, also the results of the overview were consulted.

5.3. The model – findings and synthesis of the model

This model of guideline implementation comprised six steps, successively addressing: condition characteristics; guideline characteristics; influential people; organisational factors; implementation strategies; and adherence monitoring. The steps closely followed the themes identified from the primary analysis of qualitative data (Table 4.2). The differences were that Themes 1-3 (credibility of content, credibility of source, and presentation) were grouped under 'guideline characteristics' here, and the final step 'adherence monitoring' was extracted from 'implementation strategies' theme (Figure 5.1).

5.3.1. Choosing the condition

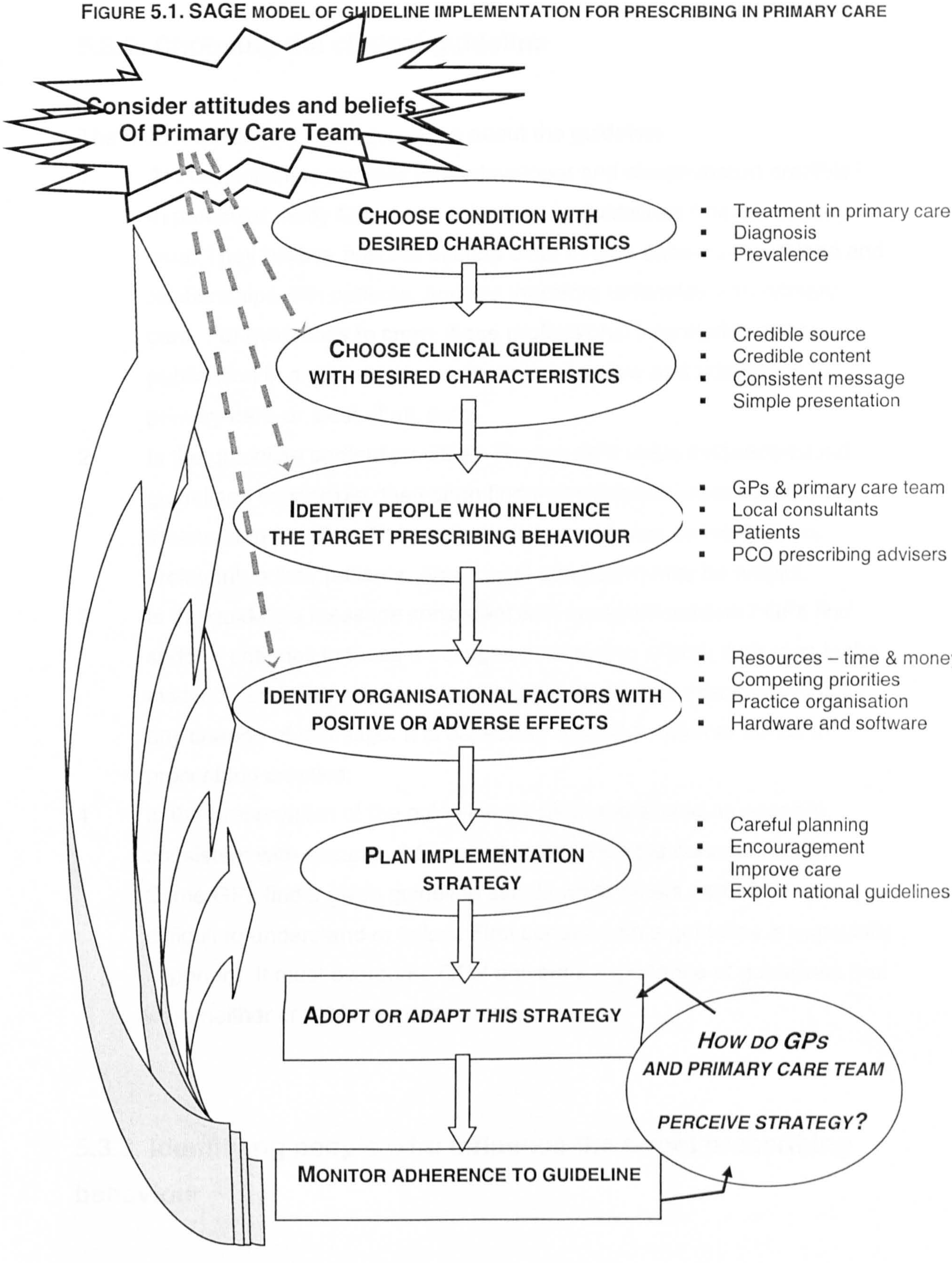
The model asked three main questions about the condition:

1 Does secondary care have a major influence on *treatment* for this condition? If so then successful implementation will need close collaboration with secondary care.

2 Why is prescribing for this condition difficult? Is it because *diagnosis* is also difficult? If so then successful implementation will need to address this issue, possibly through close collaboration with secondary care.

3 *Prevalence* – is the condition rare? If so then successful implementation needs effective reminders to GPs, perhaps through patients or computerised guidelines.

FIGURE 5.1. SAGE MODEL OF GUIDELINE IMPLEMENTATION FOR PRESCRIBING IN PRIMARY CARE



5.3.2. Choosing the clinical guideline

The model asked four main questions about the guideline:

1 Are the guideline *sources* (both developer and disseminator) credible? In particular many GPs have objections to guidelines from secondary care. They believe hospital doctors differ in their case-mix, workload and relationships with patients, and are therefore unfamiliar with primary care. For guidelines to cross these professional boundaries needs publication in a respected source, or local review and adaptation to primary care or, best of all, both.

2 Is the guideline *content* credible? Though GPs value evidence-based guidelines in principle, they often find that evidence unsatisfactory or irrelevant in practice. They may even see evidence as reliable, but irrelevant to their patients. Again local adaptation may be helpful.

3 Is the guideline *message* consistent with accepted wisdom? GPs find sudden changes in these messages threatening to both their trust in the evidence and their relationship with patients. So it is essential to justify any change of message. It is especially difficult to reverse trends in prescribing practice.

4 Is the *presentation* of the guideline as clear and simple as possible, consistent with accuracy? Simplicity is not a simple concept (4.3.5). Some GPs find a given guideline simple while others perceive it as difficult to understand or follow. First contact with a guideline is especially important. It must overcome GPs' universal experience of guidelines that were neither credible nor consistent.

5.3.3. Identifying people who influence the target prescribing behaviour

GPs within a practice generally affect each others' behaviour. The effect in larger practices can be substantial, especially when principals divide responsibilities. To ensure that one engages all opinion leaders it is prudent to

involve more than one practitioner per practice in implementation, but its benefits should be compared with extra costs involved. Other potentially influential people include practice nurses and other members of the primary care team, especially those involved in the care of patients covered by the guideline.

Outside the practice but within the locality, secondary care is a major source of influence on prescribing. It is especially important to attract local consultants' support for targeted guidelines when their existing influence is not entirely consistent with guideline recommendations. Nurse practitioners and clinical nurse specialists can also be influential. Though the influence of PCO prescribing advisers (and pharmacists) is still limited, their role in implementation is likely to increase as PCOs settle down.

Patients also influence GPs' prescribing behaviour. So targeting patients, especially those with rare diseases or major health care needs, can encourage GPs to adhere to guidelines. This can also help to maintain the patient-doctor relationship in the face of changes in guideline recommendations.

National bodies are also influential. Respondents admitted that NSFs and guidance from NICE both influence their prescribing. The main reasons given were the credibility of the sources and the authority conveyed by their publications. Pharmaceutical company representatives can also encourage or discourage adherence to prescribing guidelines. They have direct influence on GPs and indirect influence through consultants and nurses. Their role can be beneficial where the guideline recommends a new drug or more of an old drug.

5.3.4. Identifying relevant organisational factors

Respondents argued that the current state of primary care had direct effects on implementation and indirect effects on other elements in the model. Workload in general practice was a major issue. There was a strong belief that GPs and primary care teams were under increasing pressure, mainly because of continuous change in clinical practice and management.

The process of assimilating, implementing and auditing guidelines consumes yet more time. Furthermore the stepwise approach of most guidelines for drug treatment results in more appointments. Thus most GPs perceived guidelines as sources of extra workload. Fortunately GPs change their priorities to create more time for essentials. For example 'mini-clinics' have proved helpful in providing extra time and staff for targeted activities, and could help in implementing guidelines.

The cost of drugs is also important. PCOs need to be especially vigilant when advocating guidelines that recommend expensive new drugs without reducing the need for old treatments. For example the perceived cost of statins delayed their adoption until the true cost became clear. Respondents generally regarded computerised patient databases as essential in chronic disease management. Though there was more suspicion of computerised guidelines, many regarded them as useful for rare diseases or conditions with limited choices of treatment. They also observed that prescribing recommendations could be difficult to follow if access to technology were difficult, for example to echocardiography for heart failure.

5.3.5. Planning implementation strategies

PCOs should plan their strategies for implementing prescribing guidelines using the best available evidence. Many high quality evidence-based guidelines have failed through poor implementation, while some mediocre guidelines have achieved success through careful implementation strategies (Grimshaw & Russell, 1993). Therefore it is recommended that PCOs estimate the level of support needed with care. While underestimating the needs of practices will prevent successful implementation, overestimation will take resources from other guidelines.

Most of the respondents perceive central guidelines from NICE and central targets in NSFs as influential, but not yet coercive. The analysis implied

that it was important to avoid negative feelings whenever possible. The long-term effects of quality improvement programmes depend on the genuine support of those responsible for their delivery. Once GPs perceive influence as coercion, adherence to guidelines may fall. As a strategy, therefore, one should confine the enforcement of guidelines to crucial issues when other approaches have proved ineffective. Even then clear messages supported by appropriate quality indicators are essential.

The participants agreed that the main aim of prescribing guidelines is to improve patient care. They did not accept cost containment as the sole justification for guidelines. When alternative treatments differ in cost but not effectiveness, therefore, PCOs should inform GPs accordingly. They should specify the benefits of using the more cost-effective drug in releasing resources for other patients, preferably with examples. Financial incentives for adherence to guidelines were likely to be effective, and merit serious consideration.

When there is a national validated prescribing guideline for their chosen condition, PCOs should choose it and avoid weaker alternatives. Local versions of national guidelines can address local concerns, and increase both GPs' sense of ownership and their understanding of the underlying evidence. In particular PCOs can profitably encourage practices to develop protocols for adhering to national guidelines.

5.3.6. Monitoring adherence to guidelines

Quality indicators in the form of adherence measures play an important part in achieving clinical effectiveness. Their role is to identify, not only poor practice, but also outstanding improvement. Even where there is no financial reward, recognition of achievement can be fulfilling. Many practices are not aware of their strengths and adherence measures can reveal them. In short PCOs should set minimum standards of care using guidelines as yardsticks, support their practices to achieve them, and publicise success stories both locally and nationally.

5.4. Discussion

5.4.1. Taking account of GPs attitudes

Six key steps were described for PCOs and practices to consider in implementing guidelines – condition and guideline characteristics, influential people, organisational factors, and implementation and monitoring strategies. Underlying theories (Conner and Sparks, 1996) and the findings identified that practitioners' attitudes played an important role. GPs' attitudes and beliefs can moderate the influences of other factors on the effectiveness of prescribing guidelines. While those attitudes often reflect the reality of clinical practice, they can also reflect personal bias (Raisch, 1990b). Either way the attitudes of GPs and other primary care workers deserve careful attention.

5.4.2. How to use the model

The model views clinical guidelines as comprehensive tools conveying evidence to improve prescribing. The model strongly implied that PCOs should not use their limited resources to develop their own guidelines. Instead they should choose the best available guideline, preferably national, certainly validated. The membership of 'guideline teams' responsible for local adaptation, dissemination and implementation is important. They may include a wide range of stakeholders including GPs, practice and community nurses, practice managers, local consultants, community pharmacists, and clinical effectiveness officers. Team development takes time but may bring further benefits.

Guideline teams should tackle implementation as a continuous systematic cycle of quality improvement rather than a single task (Berwick, 1992; Russell and Wilson, 1992; Russell et al, 1993; Donabedian, 2003). The effectiveness of quality improvement cycles is yet to be established in critical appraisal (2.3.13), but their use seems logical and while there are not many

effective alternatives can be helpful. Some respondents recognised that clinical governance provided a useful framework for this process. Thus the co-operation and support of colleagues over time is essential. While the role of guideline characteristics and credibility receives a lot of emphasis, the findings of primary analysis (4.4.2) and other literature confirmed that this was only one of many influences on adherence (Freeman and Sweeney, 2001). Though the model gives weight to organisational factors, guideline teams should bypass those that are not alterable (Raisch, 1990a). They should also recognise that it is not cost-effective to address every shortfall in clinical behaviour (Mason et al, 2001).

5.4.3. Areas for further research

The continuing systematic review of clinical guidelines by the Cochrane Effective Practice and Organisation of Care Review Group (2002) has identified research into many elements of the model. However the contribution of other elements in the model needs more research. These include the contribution of practice nurses to GPs' prescribing behaviour. There is also a need to test empirically the ability of the theories of behaviour change summarised in this paper to explain the uptake of guidelines. In particular national sample surveys of GPs are used to validate the TPB (Ajzen, 1991) in this field (Chapters 6 and 7).

5.4.4. Summary

Changing prescribing behaviour in primary care is a difficult task to which clinical guidelines can contribute. The model provided a framework and guide for PCOs keen to improve quality through guidelines. Successful implementation depends on careful planning and consideration of the elements of the model. The model can also identify barriers that hinder adherence to guidelines. It may help to explain why clinical guidelines vary in their uptake.

The most important maxim for data analysis to heed, and one which statisticians have shunned, is this: far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.

Tukey, 1962, pp 13-14

Chapter 6: Sampling and sample size for the TPB and SAGE surveys

6.1. Introduction

In this chapter the sampling framework and sample size of the SAGE surveys are being discussed. It starts with the methods which are applied and tested for the purpose of defining a TPB survey sample size, and it follows towards specific sample size calculations for the two randomised multi-stage surveys planned for this study. The surveys' results are presented in Chapter 7. This sample size calculation work was first carried out in October 2001 prior to the main SAGE surveys. Its results were used in sampling and sample size calculation for the surveys and were presented in the International Society for Technology Assessment in Health Care (ISTAHC), Berlin (Rashidian et al, 2002). The chapter was updated in September 2003 to include new literature and to prepare it for publication. To our knowledge this is the first systematic attempt to determine sample size for TPB studies.

6.1.1. Sample size calculations in the TPB studies

Regression analyses are the main recommended analyses for the TPB (Ajzen, 1991; Conner and Sparks, 1996). Until recently most of the literature suggested simple rules-of-thumb for sample size in regression analyses (e.g. Altman, 1991, p 349). These rules-of-thumb usually take one of three forms: a minimum constant value; a minimum ratio of subjects to explanatory variables; or both combined (Green, 1991). There is a wide range of alternatives within these rules. The minimum ratio of subjects to explanatory variables, for example, varies from ten (Altman, 1991) or fifteen subjects per variable (Stevens, 1996) to forty subjects per variable in stepwise regression (Pallant, 2001). The danger of using simple rules-of-thumb is that the investigator ignores '*the idiosyncratic characteristics of research studies*' (Green, 1991, p 501). In recent years investigators have been more inclined to perform sophisticated analyses to ensure their proposed sample sizes conform to the required power. Also methodological papers are published on power in regression analysis (Green, 1991; Hsieh et al, 1998). Comparing power analysis with simple rules Green (1991) concluded '*researchers who use rules-of-thumb rather than power analysis are trading simplicity of use for accuracy and specificity of response*' (p 508).

In planning the surveys, published papers from studies of the application of the TPB and theory of reasoned action to health professionals' performance were reviewed. Out of ten papers, five reported how they estimated the required sample sizes (Table 6.1). Three papers referred to simple rules to justify their intended sample size (Levin, 1999; Meyer, 2002; Puffer and Rashidian, 2004). The fourth paper, a theory of reasoned action study (Lambert et al, 1997), used Cohen's (1988) power table and concluded a sample of only 19 would suffice. Finally, a recent randomised trial, used a computer software package for sample size calculation (Bonetti et al, 2003). A further review of ten different TPB studies (eleven papers) revealed a similar phenomenon (Anderson et al, 1998; Jemmott et al, 1998; Cox et al, 1998; Abraham et al, 1999; Sheeran et al, 1999; Conner and Mcmillan, 1999; Terry et al, 1999; Armitage and Conner, 1999a; Armitage and Conner, 1999b; Conner et al, 2000; Conner et al, 2001).

Only one of those papers reported its methods of sample size calculation (Jemmott et al, 1998). There is a wide range of sample sizes in those studies. As the minimum, a theory of reasoned action study was based on a sample of only 27 (Lambert et al, 1997) and a TPB study included 765 professionals (Millstein, 1996).

TABLE 6.1. SAMPLE SIZES IN TPB AND THEORY OF REASONED ACTION STUDIES OF HEALTH CARE PROFESSIONAL BEHAVIOURS.

Citation details	Method	Topic	Sample	Power calculation method reported?
Taylor et al, 1994	Expanded TRA; Survey	Screening mammography	85 general internists	No
Millstein, 1996	Longitudinal surveys	Offering STD prevention advice	765 doctors	No
Lambert et al, 1997	TRA; survey	Antibiotics prescribing	27 family doctors	Cohen's power table
Levin, 1999	Survey	Glove use	527 nurses and lab workers	rule of thumb for SEM
Walker et al, 2001	Survey	Antibiotic prescribing	126 GPs	No
Watson and Myers, 2001	Survey	Glove use	103 nurses	No
McCarty et al, 2001	Survey	Smoking cessation	397 nurses in for hospitals	No
Meyer, 2002	Survey	Asking for assignments	92 nurse students	N>=50+8m
Puffer and Rashidian, 2004	Survey	Offering smoking cessation advice	48 nurses	simple rule of thumb
Bonetti et al, 2003	RCT	Third molar management	99 dentists	GPower software package

TRA: the theory of reasoned action; SEM: structural equation modelling; STD: sexually transmitted diseases

Sample size calculation cannot guarantee that a study would have sufficient power (Vickers, 2003); however, where there is no power calculation it is more likely to waste resources in over-sampled or under-powered research. Given the wide range of observed sample sizes, a careful examination of power for TBP studies was considered useful for the SAGE surveys and beneficial to future research investigations. Before the exercise, two loose limits for the sample size were set: the sample should not be smaller than fifty (approximately fifteen cases per variable, Stevens, 1996). Upper limit was

considered as 600 (or two hundred per independent variable) as a larger sample could cause serious resource implications for the surveys.

6.1.2. Introducing TPB regression models

Two sets of linear regression models form the main analyses of a TPB study (Ajzen, 1991; Conner and Sparks, 1996). Model A is intended to explain the variance in behavioural intention (BI). In this model, attitudes (ATT), subjective norms (SN) and perceived behavioural controls (PBC) are the explanatory variables.

(Model A) $$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$$

or

$$BI = \beta_0 + \beta_1 \times ATT + \beta_2 \times SN + \beta_3 \times PBC$$

The second set of models (Model B) is intended to explain the variance in target behaviour (BEH). In this multiple regression BI and PBC are the explanatory variables.

(Model B) $$Y = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + e$$

or

$$BEH = \gamma_0 + \gamma_1 \times BI + \gamma_2 \times PBC$$

6.2. Methods

In all the following analyses 5% chance of Type I error was accepted, and power to detect differences between alternative hypotheses was set at 80% as recommended for research in behavioural science (Cohen, 1988).

6.2.1. Green (1991) and Cohen (1988) approach: the λ method

Green (1991) introduced a two-step approach for sample size calculation called the λ method. It is derived from previous work of Cohen (1988, pp 444-5). The method has clear advantages over other simple rules. It closely follows the power analysis approach and is capable of taking effect size into account; something that is missing in other simple rules. Green (1991) compared the output of the method with Cohen's power table and found that the results were very close, especially when there were limited numbers of explanatory variables in the regression models.

The λ method has some limitations. It is based on Cohen's (1998, pp 448-55) power tables that may be different from other power tables. It also assumes that all explanatory variables are included in the model at the same time, precluding stepwise regression analysis. This, however, is not a serious limitation for two reasons: many have recommended against the use of stepwise analysis (Miles and Shevlin, 2001; Cohen et al, 2003); and theory based analyses such as TPB studies should include all the variables into the regression model at the same time. The third limitation is the need for careful estimation of the effect sizes from prior research before the start of sample size calculation. This limitation equally applies to other methods as they too require some measure estimation from previous studies.

The measure of effect size in multiple linear regression is ρ^2 (squared multiple correlation coefficient), estimated by R^2. In order to facilitate the ease of computation, R^2 is frequently converted to f^2:

(I) Effect size index: $$f^2 = \frac{R^2}{1-R^2}$$ (Cohen, 1988, p 410)

In the λ method the null hypothesis H_0: $\rho^2 = 0$ is tested against the alternative of H_1: $\rho^2 = R^2$. Sample size (N) is a function of λ and f^2:

(II) $$N = \lambda / f^2, \text{ where } \lambda = 6.4 + 1.65m - 0.05m^2$$

m is the number of explanatory variables in the regression model. Sample size estimated by this method will be close to what can be obtained from power tables (Green, 1991).

6.2.2. Hsieh et al (1998) approach: the VIF method

For the sake of simplicity the method introduced by Hsieh et al (1998) was called the *Variance Inflation Factor (VIF)* method. As the first step in this method the sample size is calculated for a simple regression model. As the second step the sample size is adjusted for multiple regression using the estimated value of VIF (Miles and Shevlin, 2001). As the aim is to calculate the sample size for models A and B (6.1.2.), simple regression models A′ (e.g. $BI = \beta_0 + \beta_1 \times SN$) and B′ (e.g. $BEH = \gamma_0 + \gamma_1 \times BI$) are introduced to correspond with models A and B respectively. In model A′ explanatory variable can be ATT, SN or PBC depending on which one is the main focus of the research. Likewise in Model B′, BI or PBC can be used as the explanatory variable.

For the simple regression models (e.g. A′) the hypothesis H_0: $\rho = 0$ is tested against the alternative H_1: $\rho = r$, where r is the expected correlation coefficient between explanatory and outcome variables. As the models are linear, this is equivalent to testing H_0: $\beta_1 = 0$ against $\beta_1 = \beta'_1$, where β_1 is the

slope coefficient (Hsieh et al, 1998). This hypothesis tests the proposed formulae to calculate sample size (first step):

$$\text{(III)} \qquad N' = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2}{C_r^2} + 3 \qquad \text{where } C_r = \tfrac{1}{2}\log\left(\frac{1+r}{1-r}\right).$$

As $\alpha = 0.05$ and $\beta = 0.8$, hence $(Z_{1-\alpha/2} + Z_{1-\beta})^2 = (1.96 + 0.848)^2 = 7.885$

The second step is to adjust N' to correspond with the multiple regression models (A or B) for which we want to calculate the sample size. In the second step, the null hypothesis H_0: $[\beta_1, \beta_2, \beta_3] = [0, \beta_2, \beta_3]$ is tested against the alternative of $[\beta'_1, \beta_2, \beta_3]$ (Hsieh et al, 1998). To estimate the required sample size (N) for these models, an estimate of VIF and ρ^2_{123} is required. ρ^2_{123}, or tolerance, is the proportion of variance in one explanatory variable being explained by other explanatory variables in a multiple regression (V). The required sample size is:

$$\text{(IV)} \qquad N = N' \times VIF = \frac{N'}{1-\rho^2_{123}}, \text{ where } \quad VIF = \frac{1}{1-\rho^2_{123}}, \text{ and}$$

$$\text{(V)} \qquad \rho^2_{123} = \frac{r^2_{12} + r^2_{13} - 2r_{12}r_{13}r_{23}}{1-r^2_{23}}$$

Formula (V) presented in here applies to a regression model with three explanatory variables (Model A). The formula should be adjusted for the number of variables in regression model. The correlation value between the outcome and explanatory variables (r in formula III) is required for the estimation of N'. The estimation of ρ^2_{123} is possible if the zero-order correlations between explanatory variables in models A and B are available, which in turn will be used for sample size calculation (IV). The assumptions and formulae presented here can easily be modified to apply to other regression models with different number of explanatory variables.

6.2.3. Estimating the sample size

The likely values required for the SAGE surveys' sample size calculation were estimated using the reported values in published studies. Seven TPB studies of health professionals' report at least some of the values required for sample size calculation using the λ and VIF methods (Millstein, 1996; Levin, 1999; Walker et al, 2001; Watson and Myers, 2001; McCarty et al, 2001; Meyer, 2002; Puffer and Rashidian, 2004). Zero-order correlations between the TPB variables and R^2 values were extracted from studies wherever reported (Table 6.2). Then sample sizes for different models were estimated using the two methods and the results were compared to identify the appropriate method for sample size calculation. To calculate the sample size for the surveys, the median values of the reported estimates in previous studies were extracted and the appropriate method was used to calculate the basic sample size.

6.2.4. Sampling frame, design effect and stratification factor

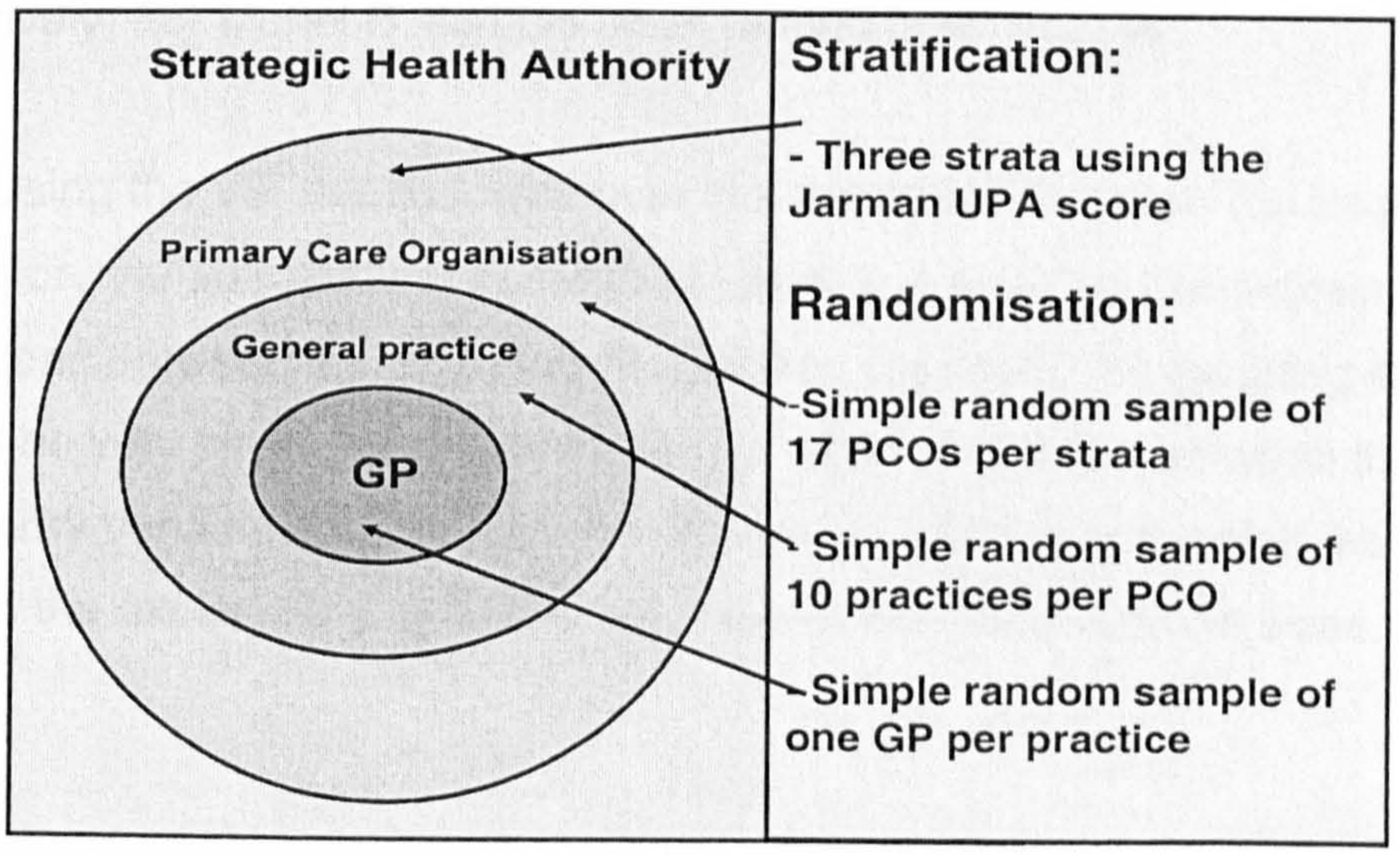
An additional issue in studies of health professionals is that the data are likely to be clustered. This means that the independence assumption is violated, and thus the power of the study is reduced. GPs who work closely together (for example in multi-partner practices or within PCOs) are likely to think and behave more similarly to one another than other GPs. For example they may have more similar prescribing patterns. Therefore the variation in prescribing of GPs working within a PCO is likely to be less than the variation observed between GPs working in different PCOs. Intra-cluster coefficient (ICC) is a statistical measure used to capture the cluster effect. In its simplest form, it can be represented as ratio of between-cluster variation to total variation (Campbell et al, 2001b): $ICC = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2}$, where σ_b^2 is between-cluster variance and σ_w^2 is within cluster variance. Bigger values of ICC represent further losses in power.

Clustering reduces effective sample size. Because of clustering effect, a larger sample size would be required to achieve the same power that could have been obtained if data were not clustered. The ratio of the total number of

subjects required using cluster randomisation to the number required using individual randomisation is known as the design effect (e.g. see Kerry and Bland, 1998). The design effect (DE) in a cluster randomised study equals to $DE = 1 + (n-1) \times ICC$ (Campbell et al, 2000b), where n is the number of cases sampled per cluster. This is accurate when cluster size is constant for all clusters and there is no stratification. Estimation of ICC is an important step in estimation of design effect. In this study ICC was estimated from published work while considering the nature of variability in prescribing for asthma and of statins. Ideally previous studies used for estimating ICC should have had similar outcome measures and have been conducted in similar setting. Given there were no studies with these characteristics sensitivity analysis was performed to establish the effects of different values of ICC on the sample size and on the optimal number of sampled GPs per PCO. A range of possible stratification variables were considered to reduce the design effect resulting from the clustered nature of primary care and to reduce random error (Moser and Kalton, 1993; Ukoumunne et al, 1999a; Ukoumunne et al, 1999b).

Figure 6.1 is a schematic presentation of the sampling frame and the randomisation process for the surveys. Further details are presented in the next Chapter (7.2.2).

FIGURE 6.1. SAGE MULTI-STAGE SAMPLING SCHEME



6.3. Findings

6.3.1. Sample size by the λ method

Sample size was calculated using the reported R^2 values (Table 6.2.). The resulting sample sizes ranged from four to 41 for Model A (prediction of intention) and from five to ten for Model C (prediction of behaviour). These sample sizes were smaller than the minimum samples suggested by most rules-of-thumb. Small sample sizes were the outcomes of large reported R^2 values of up to 0.74 in previous TPB studies (Table 6.2).

6.3.2. Sample size by the VIF method

This method requires the estimated values of zero-order correlations for sample size calculation. Many papers do not report the zero-order correlation values between explanatory variables. In this study, sample size calculation with this method provided a wide range of values and larger than what achieved through the λ method (Table 6.2). For Model A, it ranged from 36 to 4024, resulting from correlation coefficients of 0.56 and 0.06 between subjective norm and intention respectively. For Model B, sample sizes ranged from 18 to 92.

Using the VIF method, wherever the correlation between the target explanatory variable and the dependent variable is large and zero-order correlations between the explanatory variables are small, the resulting sample size will be very small. On the other hand a small correlation between the target explanatory variable and the dependent variable, with large correlations between the explanatory variables will result in very large sample sizes.

TABLE 6.2. REPORTED CORRELATIONS BETWEEN TPB VARIABLES AND RESULTING SAMPLE SIZES USING THE VIF [F] AND λ METHODS

Authors	Sample size	Model R^2 [a]		Correlations						N for Model A		N for Model B	
		Model A	*Model B*	*BI – SN*	*ATT - SN*	*ATT - PBC*	*SN - PBC*	*BEH - BI*	*BI - PBC*	*VIF* (SN in simple model)	λ	*VIF* (BI in simple model)	λ
Millstein, 1996	765	0.27	0.39 [b]	N/R	N/R	N/R	N/R	N/R	N/R		29		
Levin, 1999	527	0.74	0.66 [c]	0.28 0.34	0.54	0.32 0.46	0.23 0.34	0.35 0.44	0.43 0.57	140 95 146 [d]	4	76 57 92 [d]	5
McCarty et al, 2001	397	~0.23 [e]	N/A	0.30	0.41	0.46	0.46	N/A	0.42	116	35		
Walker et al, 2001	126	0.48	N/A	0.38	0.49	0.28	0.21	N/A	0.08 (ns)	70	12		
Watson and Myers, 2001	103	0.44	0.50 [c]	0.07 (ns)	0.13 (ns)	0.37	0.08 (ns)	0.69	0.43	1637	14	18	10
Meyer, 2002	93	~0.48 [e]	N/A	0.56	0.56	0.49	0.46	N/A	0.43	36	12		
Puffer and Rashidian, 2004	48	0.21 0.40	N/A	0.06 (ns)	0.67	0.17 (ns)	0.03 (ns)	N/A	0.55	4024	41 17		

Model A: $BI = \beta_0 + \beta_1 \times ATT + \beta_2 \times SN + \beta_3 \times PBC$; Model B: $BEH = \gamma_0 + \gamma_1 \times BI + \gamma_2 \times PBC$; ns: not significant; N/A: not applicable; N/R: not reported; [a] values corresponding to those models (A and B) are reported. These are not necessarily the highest value reported in the paper or the abstract; [b] full set of variables; [c] BEH measured at the same time as BI; [d] Maximum sample size calculated using the correlations provided; [e] Estimated using correlations values; [f] In this table we only introduce the sample sizes calculated for Models A' and B' as introduced in the paper. We have performed further analyses for alternative simple regression models. The findings of those analyses are available from the authors upon request.

6.3.3. SAGE surveys' basic sample sizes

The analyses demonstrated that within the range of likely values reported in TPB studies, the VIF was more sensitive to variation in parameters' estimates than the λ method (Table 6.2). The VIF was chosen as the method of choice for two reasons. First, TPB studies tended to report high goodness-of-fit values. Therefore, sample size calculation methods that used the goodness-of-fit of the model would result in sample sizes estimates that were too small for stable regression analyses. Second, the λ method ignored the peculiarities of potential collinearities between the TPB explanatory variables. It also ignored variations in the correlations between the explanatory variables and the independent variables. Therefore, sample sizes estimated by the λ method might be too small for identification of important contributions made by individual variables into the model. One might argue that in TPB studies the contribution of individual explanatory variables in the models is more informative than the general goodness-of-fit of the models.

Thus the VIF method was used for the SAGE surveys sample size calculation. To calculate the surveys' sample sizes, the median values of the reported coefficients in previous studies summarised in Table 6.2 were used (Table 6.3). The only exception was the correlation between intention and behaviour, where the median value of 0.4 from previous studies was not used. Instead the sample size was calculated in a way to have power of detecting a correlation coefficient of 0.25 between intention and behaviour. This was due to the reported correlation coefficients were from two studies that measured 'self-reported' behaviour at the same time of measuring intention resulting in artificially high reported coefficients and R^2 values (Levin, 1999; Watson and Myers, 2001). The sample sizes were estimated to be 114 for Model A, and 148 for Model B, if simple random sampling was used (Table 6.3).

TABLE 6.3. SAGE SAMPLE SIZE, BEFORE INCORPORATION OF THE DESIGN EFFECT

Model A				
Correlations	SN and BI	SN and ATT	SN and PBC	ATT and PBC
	0.3	0.5	0.2	0.4
Sample size	**114**			
Model C				
Correlations	BI and BEH	BI and PBC		
	0.25	0.4		
Sample size	**148**			

6.3.4. ICC and design effect

The outcome variables of interest in here are GPs' prescribing for asthma and of statins, both process variables (Campbell et al, 2000b). The variables should be adjusted for case-mix and demographic variations, before being entered in the analysis. Many other patient and practitioner variables also affect prescribing. Doctors working in one locality may prescribe more similarly than those who work in different areas, hence increasing ICC. If prescribing data is collected using GP records, then it is a process variable which heavily relies on clinician and practice characteristics. It is also influenced somehow by patients, as patients' characteristics influence prescribing patterns. On the other hand prescribing data obtained from dispensing sources or pharmacies is a process variable that becomes more correlated with patients' compliance and characteristics. Dispensing data may be different from what can be obtained from general practices, as some patients may decide not to take their prescriptions to the dispenser. This suggests prescribing data from pharmacies has the potential of more variability within clusters than it is the case for process variables. Prescribing data to be used for the analysis of the surveys was obtained through access to the Prescribing Analyses and CosT (PACT) data (Majeed et al, 1997).

Campbell et al (2000b) estimated that ICCs for process variables were in range of 0.05-0.15. They suggested using ICC value of 0.1 for process variables where there was no reliable estimate. ICCs tend to decrease in bigger clusters (Ukoumunne et al, 1999a). In the SAGE surveys, GPs were clustered within PCOs. The arguments put forward in the previous paragraph and the size of the clusters suggested that the true value of ICC for the SAGE surveys was probably smaller than 0.1. There were no published ICCs measures for the process variables to be used in the surveys (i.e. prescribing for asthma and of statins). To decide the likely value of the ICCs within the range of 0.05-0.1 it was decided to look for the ICCs for similar patient outcomes. Ukoumunne et al (1999, pp 74-75) reported ICCs of 0.0096 for 'taking medication for asthma', 0.0000 for 'having steroid inhaler', 0.0061 for 'having asthma not on inhaled beta-2 agonists' and 0.0011 for 'having asthma not on inhaled steroids' from two different British studies in primary care. Although these variables were patient outcomes clustered within practices, they were the closest variables to asthma prescribing with reported ICCs that were identified in the literature. These small values implied that ICCs for different measures of asthma prescribing might be close to the minimum value suggested by Campbell et al (2000b).

ICC for statins may be slightly higher than asthma. The qualitative study (Chapter 4) suggested that statins prescribing was more sensitive to prescribing costs and budget limitations. Also local initiatives might have greater influence on statins prescribing especially for primary prevention of coronary heart disease (Rashidian and Russell, 2003). More similarities were observed between the interviewees in their intention to prescribe preventive drugs for asthma and to follow the recommendations of the guidelines and more variability in their views towards statins guidelines.

Thus it seemed that an ICC of 0.1 to be unnecessarily large for the SAGE surveys and a value closer to 0.05 was more representative of the true ICC. Although the ICC values for asthma and statins prescribing were not necessarily the same, using a single value for both eased the sampling procedures and the surveys' handling. With all of these considerations, a value of 0.07 for ICC seemed reasonable. Sensitivity analysis of the effects of

different ICC values and cluster sizes were performed for both sample sizes of 114 (Table 6.4) and 148 (Table 6.5). The tables' format was derived from Ukoumunne et al (1999a). With an ICC of 0.07 and 10 GPs per cluster, the required sample size for Model B was 242.

TABLE 6.4. SENSITIVITY ANALYSIS FOR THE EFFECTS OF ICC AND CLUSTER SIZE ON DESIGN EFFECT AND SAMPLE SIZE, IF THE SIMPLE RANDOM SAMPLE SIZE IS 114

Estimated ICC	Number sampled per cluster $n = 10$			Number sampled per cluster $n = 20$		
	Design effect	GPs required	Clusters required	Design effect	GPs required	Clusters required
0.00	1.00	114	12	1.00	114	6
0.01	1.09	125	13	1.19	136	7
0.03	1.27	145	15	1.57	179	9
0.05	1.45	166	17	1.95	223	12
0.07	1.63	**186**	19	2.33	**266**	14
0.10	1.9	217	22	2.9	331	17

TABLE 6.5. SENSITIVITY ANALYSIS FOR THE EFFECTS OF ICC AND CLUSTER SIZE ON DESIGN EFFECT AND SAMPLE SIZE, IF THE SIMPLE RANDOM SAMPLE SIZE IS 148

Estimated ICC	Number sampled per cluster $n = 10$			Number sampled per cluster $n = 20$		
	Design effect	GPs required	Clusters required	Design effect	GPs required	Clusters required
0.00	1.00	148	15	1.00	148	8
0.01	1.09	162	17	1.19	177	9
0.03	1.27	188	19	1.57	233	12
0.05	1.45	215	22	1.95	289	15
0.07	1.63	**242**	25	2.33	**345**	18
0.10	1.9	282	29	2.9	430	22

6.3.5. Stratification factor

Stratification reduces the design effect and randomisation errors (Moser and Kalton, 1993; Robson, 1993; Ukoumunne et al, 1999a). The measures of the target behaviour (dependent variables) are the most appropriate variables for stratification. The target behaviours in here were prescribing for asthma and

statins as documented in PACT data. Neither of these outcome measures were available prior to the surveys. Therefore a range of other variables were considered for stratification. The Jarman Underprivileged Area score (Jarman score) was used for stratification (Jarman, 1983).

6.3.6. Response rate and the final sample size

Response rates to surveys including surveys of GPs varies considerably (Asch et al, 1997; Olatunbosun et al, 1998; Di Iorio et al, 2000; McColl et al, 2001). By employing different approaches for increasing the response rate (including individually signed covering letters, information sheets, good quality prints, reminder letters, business reply envelopes etc (McColl et al, 2001; Edwards et al, 2002)) a response rate of 50% or higher deemed attainable. Considering a response rate of 50%, final proposed sample size for each survey equalled 484 (242/0.50; Figure 6.2).

FIGURE 6.2. ESTIMATED SAMPLE SIZE FOR MODEL B [A]

Simple random sampling, using VIF method ($\alpha = 0.05$; power = 80%)	148
Design Effect = 1.63 (ICC = 0.07; 10 GPs per cluster)	242
50% response rate	484

[A] $BEH = \beta_0 + \beta_1 \times BI + \beta_2 \times PBC$

6.4. Discussion

6.4.1. Method of choice for sample size calculation

The findings in Table 6.2 suggested the λ method had limited application for TPB sample size when it was based on actual goodness-of-fit (R^2) values derived from papers, rather than values suggested by Cohen (1988). Green (1991) provided evidence for the advantage of the λ method over other less sophisticated rules (e.g. N> 50 + 8m) when generic values were used for models' goodness-of-fit. Within the TPB literature the regression models' goodness-of-fit tended to be large (Table 6.2). All the reported goodness-of-fit values were larger than 0.2 with a median of 0.48. Hence the estimated sample sizes ranged from (as low as) four to 41 (Table 6.2). These sample sizes were less than minimum samples suggested by most rules-of-thumb. In this context it was concluded that the λ method had no advantage over simple rules-of-thumb and could be misleading (also see 6.3.3). An example of erroneous judgments that followed the use of the λ method was seen in Lambert et al (1997) study when they concluded a sample of 19 would suffice.

The VIF method was responsive to the requirements of TPB studies. For Model A, the power calculation was based on the least powerful predictor of intention, subjective norm. This way the study would detect significant contributions of all explanatory variables in variation in intention. For Model B, the power calculation was based on the best predictor of behaviour, intention because a small correlation coefficient between prescribing and predictors was possible. The sample size calculations resulted in larger sample sizes for Model A than for Model B (Table 6.2). The opposite was expected because the TPB variables were more powerful in explaining the variance in intention than in behaviour. Smaller sample sizes for Model B in here were merely the result of the limitation of information available. Table 6.2 demonstrated that only three TPB studies attempted to measure behaviour as well as intention (Millstein, 1996; Levin, 1999; Watson and Myers, 2001). Of those, only two papers

provided values required for the VIF method (Levin, 1999; Watson and Myers, 2001). In both studies behaviour was self-reported and measured at the time of measuring intentions. It is known that simultaneous measurement of intention and behaviour is flawed (Conner and Sparks, 1996) and erroneously inflates the effects size. Thus larger sample sizes are required for the TPB studies where variation in the behaviour is explained.

6.4.2. Estimating ICC

We could have used other approaches for the estimation of ICC. Some investigators use routine data (Eldridge et al, 2001). Use of routine data can lead to more accurate estimates of ICC, but it requires access to data sources and can be time consuming. This would have been the preferred option, given there were no previous reports of the ICCs for the study outcomes in the literature. Unfortunately the study did not have access to the routine data required for the calculation of the ICCs.

There are concerns about the reliability of ICC estimates from previous studies. Difference in outcome measures and settings limits the generalisability. Also ICC measurement is dependent on the estimation methods and hence subject to variability (Muller and Buttner, 1994; Campbell, 2000). It can also be manipulated by the choice of independent variable (Campbell, 2000). Nonetheless, obtaining ICCs from previous studies is the most widely used and a practical approach. Others tried to assign confidence intervals to the estimates (Ukoumunne, 2002). The outcomes of this work were not encouraging as the confidence intervals were very wide.

6.4.3. Choice of stratification factor

Although the Jarman score was used for stratification in the surveys, there were other variables that could have been used for this purpose. One potential approach was to use the presence of specific clinics for asthma and coronary

heart diseases in practices as a proxy of quality of care and stratify according to it. Data availability was the advantage of this approach. However, this factor could not effectively discriminate between practices. The majority of general practices provided coronary heart disease prevention special clinics and more than 90% of the general practices provided special clinics for the treatment of asthma (Baker and Hann, 2001).

Larger practices have more enhanced infra-structures and capacities including support from nurse practitioners than solo or smaller practices. It was possible to stratify based on the number of whole time equivalent GPs in practices and consider the variable as a proxy to the quality of care; because infra-structure and primary care professionals contributed to variation in the quality of care provided by GPs (incl. statins and asthma prescribing). Some recent studies challenged this hypothesis. One paper suggested single-handed practitioner did not necessarily provide lower quality of care (Hippisley-Cox et al, 2001). A large scale national survey in England concluded that although practice size was related with quality of care, this was not a simple relationship and different types of practices had different advantages (Campbell et al, 2001a). An RCT of educational outreach visits to improve prescribing for four clinical conditions concluded that single-handed practices were more likely to improve their compliance with evidence-based clinical guidelines than multi-partner practices (Freemantle et al, 2002). Another study reported that PCGs with bigger proportion of single-handed practices did not necessarily had lower provision of some services, including asthma clinics (Baker and Hann, 2001). A further study of statins prescribing concluded that the number of partners in practice and practice characteristics (fund-holding and training statuses) had no significant relationship with statins prescribing (Packham et al, 1999). The average list size per whole time equivalent GP at practitioner, practice or PCO level could also be of some value for stratification. There might be a correlation between the workload and the quality of prescribing. In this case list size should have been adjusted for sex ratio and age of the patients.

There were justifications for the use of socioeconomic variables as the stratification factor. Socioeconomic status of the patients might influence the quality of care offered to the patients. For example it was argued that practices

with a greater proportion of middle-class or younger patients were more likely to improve than others (Jefferys and Sachs, 1983; Horder et al, 1986). Deprivation scores were shown to be significantly correlated with all cause cardiovascular disease mortality and hence were valid measures of need (Smith et al, 1998b). Several deprivations scores are used for the identification of area socioeconomic status. The Jarman score, like some others, is based on census data. It is originally developed as the measure of GP workload (Jarman, 1983; Jarman, 1984). It is known that census based scores may have limited validity in representing actual values (Scrivener and Lloyd, 1995). They may also be out of date, depending on the timing of the application in comparison to the census year. Despite all, the Jarman score has shown useful applications in different settings as a socioeconomic variable. It has been validated as a measure of health care need (Foreman et al, 2003) including for asthma and statins prescribing (Packham et al, 1999; Salamzadeh et al, 2003). It is widely used in a wide range of different studies as a proxy for (health care) need and or deprivation.

The Jarman score possesses some basic advantages that improve its validity as a stratifying factor in the SAGE surveys. The score combines a series of demographic characteristics (e.g. elderly living alone, children under five in the area), socioeconomic circumstances (unemployment and social class V) and also the proportion of people from ethnic minorities (Jarman, 1983; Baker and Hann, 2001). Those characteristics may influence the need for care (asthma and coronary heart disease) and also the compliance with the offered care. A UK study measured the rate of parents' reported wheezing symptoms in children. It concluded that lower social class correlates with higher frequency of persistent wheezing in children (Duran-Tauleria and Rona, 1999). Having ethnic minorities among the indicators is another advantage of the Jarman score. There is some evidence that ethnic minorities receive lower quality of care including for asthma care (Griffiths et al, 2001), although ethnic background may not be directly related to the number of first asthma admissions to hospitals (Griffiths et al, 1997). A significant relationship between the Jarman score and prescribing for asthma, as obtained from PACT, was explored in a study of GPs in Bradford (Salamzadeh et al, 2001). They also reported a higher level of readmission to hospital in asthma patients from high Jarman score areas

(Salamzadeh et al, 2003). This study was different from Griffiths et al (1997) as it focused on readmission rate instead of first admissions, which are more linked to individual patient circumstances. Despite these, Baker and Hann (2001) found an inconsistent relationship between the Jarman score and the provision of asthma care in practices in different areas of the UK. Another study observed no significant relationship between the Jarman score and some process measures of asthma care (Campbell et al, 2001a). However, prescribing was the outcome measure in neither of the two latter studies.

Another widely used deprivation score is the Townsend score. It comprises four variables which basically represent the socioeconomic status of households: unemployment, car or home ownership and overcrowding (Townsend, 1987). This combination of items may explain the lack of relationship between this score and prescribing for asthma reported from PACT (Salamzedeh et al, 2001). On the other hand a relatively coherent relationship is observed between higher Townsend score in the area and higher prevalence of reported persistent wheezing in children (Duran-Tauleria and Rona, 1999).

There are also important relationships between some deprivation scores and statins prescribing. One study reported a significant negative relationship between the level of deprivation (measured by the Townsend or Jarman scores) and statins prescribing (Packham et al, 1999). Higher utilisation of cardiovascular treatments in low income families, compared with the social classes of I and II, is reported in the literature (Lloyd et al, 1995b). The phenomenon may be due to higher prevalence of the disease. Both the Townsend and Jarman scores identify low income families and they are highly correlated (Pearson r=0.89 reported in Lloyd et al, 1995b).

Use of census based data for stratification of general practices is prone to some biases. These data are reported according to the postal code. If the GP practice postcode is used, the problem is that patients registered with the practice may live in addresses with different postcodes and hence deprivation scores. If the scores are used at the PCO level, this problem will be less prominent but then stratification will be limited to the PCO level. A better solution is to measure the practice score through the collation of the postcodes

of all the patients' on the practice list (a weighted average). This is previously used as a measure of need to construct the prescribing allocation formula at PCG level (Rice et al, 2000). The basic unit of analysis in the Rice at al (2000) study was practice but the data can be aggregated to PCO level if their constituent practices are known (personal communication with N Rice, Sep 2001). Nonetheless, access to the collated deprivations scores was not granted to this study. As a result and given the observed relationships between the Jarman score with statins and asthma prescribing, this score was used as the stratifying factor at the PCO level.

6.5. Conclusions

If theoretical implications are to be tested, for example the mediating role of intention between attitudes and behaviour, then the study should be powered for that purpose. A review of the theory of reasoned action studies concluded many studies were powerful enough to test the goodness-of-fit (R^2) of the model, but not so to test the mediating effects of intentions between attitudes and behaviour (Bagozzi et al, 1989). The study supported their finding. It could be argued that if the approach presented in this paper was followed, there should be enough power to test the models' specific elements.

The relationships of interest within the regression models of the SAGE surveys are between subjective norm and intention (Model A), and between intention and behaviour (Model B). Given a generally weak relationship between subjective norm and intention the proposed sample size will be large enough to reject the null hypothesis of no correlation between either attitude and intention or behavioural control and intention (Armitage and Conner, 2001). However, as the correlations between behavioural control and behaviour are usually smaller than those between intention and behaviour, the study is likely to be underpowered for the weaker elements in Model B. For both models, the proposed sample size is indeed large enough to test the significance of the models as a whole.

To our knowledge this was the first comprehensive attempt to establish a suitable method for sample size calculation for a TPB study. Most of the published studies of the TPB had ignored the issue of sample size. Two methods of sample size calculation for regression analysis were compared. The findings suggested that the VIF method was the most appropriate approach, especially when the contribution of specific explanatory variables in the model was sought.

When we cannot measure, our knowledge is meagre and imperfect.

Lord Kelvin, British scientist

Even when we can measure, our knowledge is meagre and imperfect.

Jacob Viner, American economist; quoted from Fuchs, 2000, p142.

Chapter 7. Attitudes, beliefs and prescribing for asthma and of statins: national surveys of GPs

7.1. Introduction

This chapter reports the results of two surveys of GPs beliefs, attitudes, prescribing intentions and prescribing outcomes. The two surveys were used to assess the ability of the TPB in capturing observed variations in GP prescribing intention and prescribing behaviour. The analyses present the limitations and merits of TPB in this setting. The results of the surveys are then compared with the findings of the previous stages of the project.

The results of the qualitative study (Chapter 4) were used to identify suitable clinical and prescribing topics for the assessment of TPB. Secondary prevention of CHD and management of asthma were chosen according to two criteria. First, clinical conditions should have had at least one nationally known clinical guideline in order to be able to assess GPs intentions to adhere to the recommendations. Both CHD prevention and treatment of asthma had nationally known clinical guidelines. Second, the drug therapies used for those clinical conditions should have limited applications for the treatment of other diseases. Both statins and inhaled corticosteroids had limited applications for

diseases other than CHD or asthma. The latter criterion was required as PACT data was not linked to individual patient characteristics or diagnoses.

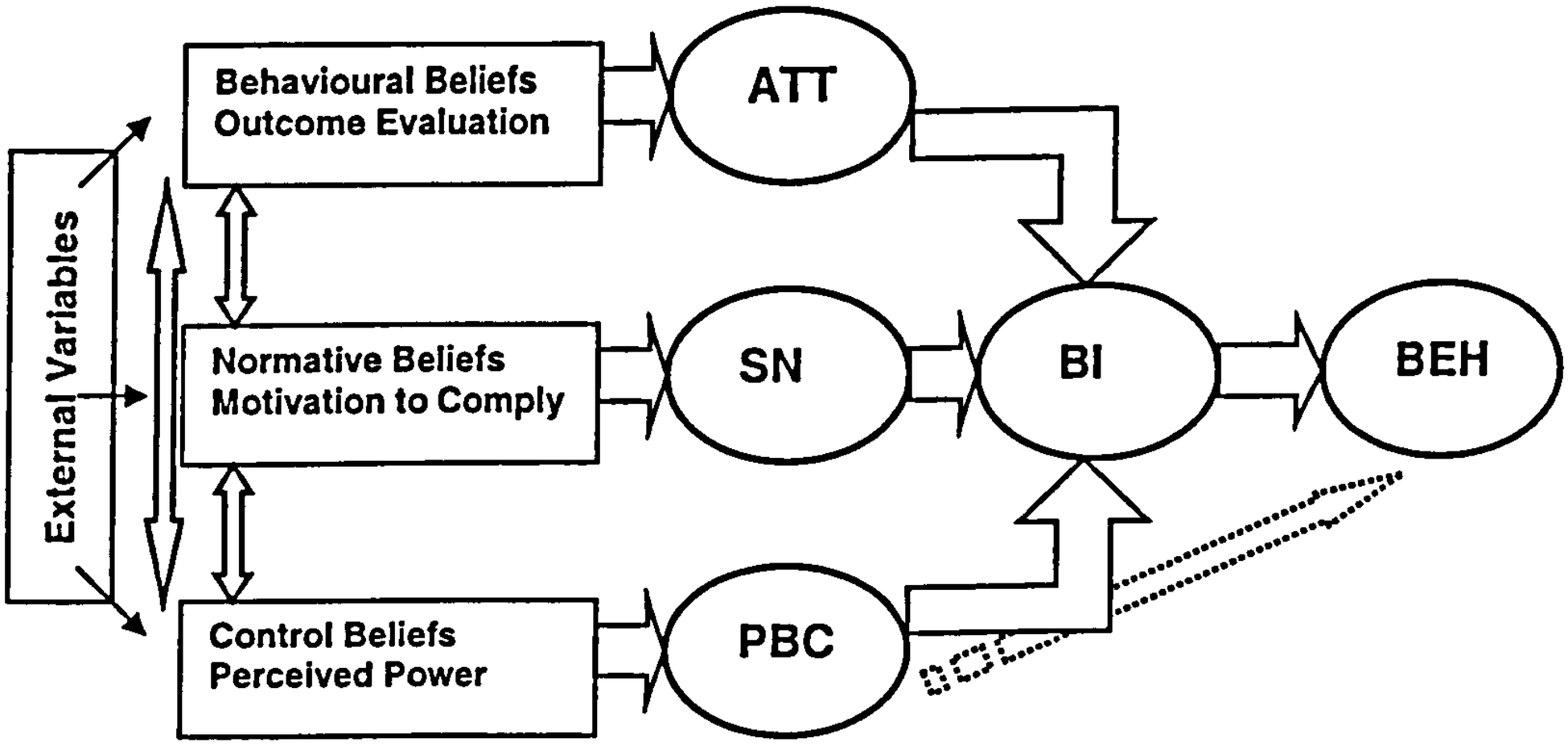
7.1.1. Theory of planned behaviour (TPB) – a summary

Chapter 3 provided more detailed presentation of TPB (3.2.6). Here TPB is briefly reviewed. TPB (Ajzen, 1991) is a social cognition theory (Conner and Sparks, 1996), extended from the theory of reasoned action (Ajzen and Fishbein, 1980) to provide a theoretical basis for explaining volitional behaviour. The theory assumes individuals are *rational actors* who carefully process information before making behavioural decisions. The information processing may change underlying beliefs and through them behaviour.

Theory of planned behaviour suggests that behavioural intention (intention) is determined by attitude towards the behaviour (attitude), perceived social or peer pressure (subjective norm) and perceived behavioural control (perceived control) (Figure 7.1). The term 'attitude towards behaviour' is to prevent confusing different attitudes with one another (Montano and Kasprzyk, 2002). For example a GP's attitude towards asthma can be different from his attitude towards prescribing a drug to treat asthma. The subjective norm is to explain the effects of perceived expectation of important others on the person's behaviour. A GP may perceive pressure from respected colleagues to prescribe for a given condition, and hence be more inclined to prescribe. An effect in the opposite direction may result from perceived pressure from someone considered untrustworthy or not-respected. Perceived behavioural control is the person's perception of his ability to perform the behaviour. According to the theory, internal and external factors might inhibit or facilitate behaviour through control beliefs (Conner and Sparks, 1996). Examples of internal factors are information, skills, emotions, and personal capabilities. Examples of external factors are opportunities, resources, and social and organisational barriers and facilitators. Theory of planned behaviour acknowledges that in many circumstances individuals perform (or refrain from performing) behaviours despite their intention. This is reflected in the model through a link between

perceived behavioural control and behaviour (Figure 7.1). Theoretically the TPB is capable of taking into account the person's past behaviour. Past experiences affect the person's perceived control and contributes to the formation of intention and behaviour.

FIGURE 7.1. THEORY OF PLANNED BEHAVIOUR (TPB)



Regression models:

Model A) $BI = \beta_0 + \beta_1 \times ATT + \beta_2 \times SN + \beta_3 \times PBC$

Model B) $BEH = \gamma_0 + \gamma_1 \times BI + \gamma_2 \times PBC$

BEH: Behaviour; **BI:** Behavioural intention; **ATT:** Attitude towards behaviour; **SN:** Subjective norm; **PBC:** Perceived behavioural control

The theory proposes linear relationships between model elements. Attitude, subjective norm and perceived behavioural control all explain intention. In turn, intention and perceived behavioural control predict behaviour. These relationships are presented by two regression models (A and B) in Figure 7.1 (Ajzen, 1991; Conner and Sparks, 1996). Theory of planned behaviour has shown a good ability in prediction of intended behaviour. The general idea is that the professional behaviour of health care professionals is mostly intentional. If it is true, the TPB could explain some of the variation in their behaviour. In particular TPB might be able to reveal some of the important barriers to and facilitators of evidence-based behaviour. For successful change of behaviour, careful attention to attitudes is one of the important factors (Oxman and Flottorp, 2001). Attitudes of physicians toward guidelines have

been studied before (e.g. Olatunbosun et al, 1998). The advantage of TPB is in its consideration of subjective norms and perceived controls.

7.1.2. What is PACT data? What are prescribing units?

Prescribing analyses and cost (PACT) counts all GP prescriptions that have been dispensed by community pharmacies (i.e. not secondary care based), dispensing practices or appliance contractors (Majeed et al, 1997). The PACT 'catalogue' includes quantity and number of items prescribed and their costs. It is difficult to assess the quality and relevance of the prescribing activities reported in PACT as the data are not linked to patients or diagnoses. PACT is widely used as a health services research tool (for example see Hobbs et al, 1996; Majeed et al, 1997; Griffiths et al, 1997; Watson et al, 2001; Jones et al, 2001; Watkins et al, 2003).

General practice demography influences GP prescribing patterns (Majeed et al, 1997). For example a university health centre is likely to have lower prescribing rates for statins on average since many registered patients are young adults. Prescribing units are therefore developed to weight age and gender to reflect for prescribing variation caused by demographic differences. Prescribing units reflect approximate prescribing needs of the population served by a practice and are based on studies of datasets linked to patients (Majeed et al, 1997). The simplest forms of prescribing units are those based on the number of patients registered with a practice, with higher weights given to elderly patients. A more sophisticated prescribing unit takes into account age and gender as well as other factors (Lloyd et al, 1997). This latter prescribing unit (known as ASTRO-PUs) is applicable across diseases as general indicator of prescribing need, hence may have limited validity when specific diseases or classes of drugs are assessed. Specific therapeutic group age-sex related prescribing units (STAR-PUs) take into account disease and treatment category and are calculated for different sections of the BNF (Lloyd et al, 1995a).

7.1.3. What are inhaled corticosteroids for?

Inhaled short-acting β_2-agonist bronchodilators (e.g. salbutamol) are the first line of treatment for mild asthma. Previous studies suggested that about half of asthmatic patients were being treated using inhaled short-acting β_2-agonist bronchodilators only (Walsh et al, 1999). Chronic and frequent use of these products is not recommended and their use should be limited to control of exacerbation of asthmatic signs and symptoms (British Thoracic Society and National Asthma Campaign, 1997a; Scottish Intercollegiate Guidelines Network, 1998; Eccles et al, 2001b); although this advice may not be followed for an important proportion of patients (Warner, 1995; Walsh et al, 1999). Inhaled corticosteroids are used as the second line of treatment. The main indication for the use of inhaled corticosteroids is for the management of asthma in patients not controlled by, or requiring frequent use of, short-acting β_2-agonist bronchodilators (British Thoracic Society and National Asthma Campaign, 1997a; Prescribing Support Unit, 2002). Inhaled corticosteroids used for prophylactic treatment of asthma are of three main types: beclomethasone dipropionate, budesonide and fluticasone propionate (British Medical Association and Royal Pharmaceutical Society of Great Britain, 2001). There are some differences between these. The most commonly used and cheapest product is beclomethasone. Budesonide has a similar efficacy to beclomethasone (Boe et al, 1989), but lower systemic absorption makes it more useful for patients requiring higher doses of steroids. Fluticasone is twice as potent as beclomethasone (Barnes et al, 1993) and also more expensive (Prescribing Support Unit, 2002). Fluticasone is more likely to be consultant initiated than the others (Anon., 2002). There are also combination preparations including inhaled corticosteroids and short- or long-acting β_2-agonist bronchodilators. Examples of these are beclomethasone & salbutamol, budesonide & formeterol and fluticasone & salmeterol. The combination products are generally more expensive than single products, but easier to administer to patients. These medicines are administered by different devices, in different dosages, and as generic or non-generic preparations affecting both their efficacy and cost. For patients requiring large quantities of corticosteroids, addition of long-acting β_2-agonist bronchodilators (e.g. formoterol or salmeterol) is recommended. These drugs are prescribed as separate formulae or in

combination with corticosteroids. It should be noted that these medicines do not replace corticosteroids, but are prescribed as addition to them to reduce the need for corticosteroids (British Thoracic Society and National Asthma Campaign, 1997a). Other groups of drugs including those of theophylline family are also used for treatment of asthma. β_2-agonist bronchodilators and theophyllines are used for other indications including chronic obstructive pulmonary disease (COPD). The use of inhaled corticosteroid is almost exclusively limited to asthma (Prescribing Support Unit, 2002). As there is no strong evidence suggesting clinical superiority of one drug over the others among inhaled corticosteroids, short-acting bronchodilators or long-acting bronchodilators, it is recommended that the prescribing of the cheapest drug in each category should be considered (Eccles et al, 2001b).

7.1.4. What are statins for?

Administration of statins for secondary prevention of CHD reduces mortality and coronary artery events. A large survey of general practices in England in 1997-1998 revealed that still many patients were not receiving the quality of care expected for them for the secondary prevention of coronary heart diseases (Brady et al, 2001). They revealed that in all areas of secondary prevention there were 'ample opportunities' for improvement. Hypercholesterolaemia was less well managed than hypertension with many patients not receiving statins that they required (Brady et al, 2001). Currently only 35% of patients who require statins are receiving them (Malik, 2004).

Statins are 3-hydroxy-3-methylglutaryl coenzyme-A (HMGCoA) reductase inhibitors, and generally result in reduction in low-density lipoprotein cholesterol (LDL) and total cholesterol levels. They have been proved effective for secondary and primary prevention of CHD in several RCTs (Caro et al, 1997; West of Scotland Coronary Prevention Study Group, 1998; Tonkin et al, 2000; Schwartz et al, 2001; Sever et al, 2003). Four types of statins were in use in the UK at the time of the study: atrovastatin, pravastatin, simvastatin and fluvastatin. A fifth type (cerivastatin) had been withdrawn from the British market

owing to its side-effects. The most commonly used statins are simvastatin and pravastatin. Atrovastatin has the added advantage of being effective for 'combined hyperlipidaemia' (i.e. high blood levels of cholesterol and triglyceride together; British Medical Association and Royal Pharmaceutical Society of Great Britain, 2001). Fluvastatin is the least potent and atrovastatin is the most potent statin (Jones et al, 1998; Prescribing Support Unit, 2002). The main indications for the use of statins are primary and secondary prevention of the CHD and also control of hypercholesterolaemia.

7.1.5. Objectives

The surveys aimed to assess the ability of TPB to explain and predict GPs' prescribing in accordance with clinical guidelines. The aim was pursued through the following objectives.

- To explore the ability of TPB to explain GPs' intentions to adhere to national guidelines on prescribing statins
- To explore the ability of TPB to explain GPs' intentions to adhere to national guidelines on prescribing for asthma
- To explore the ability of TPB to explain and predict GPs' prescribing of statins
- To explore the ability of TPB to explain and predict GPs' prescribing for asthma
- To explore practice and demographic variables that can explain GPs' prescribing of statins and for asthma
- To explore optimal scaling strategies for internal reliability and analytical qualities of TPB measures in the context of GP prescribing

7.2. Methods

7.2.1. Setting, participants and data collection

Two stratified random samples of GPs across England were surveyed. Pilot surveys were conducted in 2001. The first mailings of the main surveys were sent out towards the end of January 2002. Two reminders followed the main surveys at intervals of approximately twenty days. The last filled in questionnaires were received in early April 2002. All the mailings included the questionnaire (Appendices VII-1 or VII-2), pre-paid reply envelope, individually addressed and hand-signed covering letter (Appendix VII-4) and information sheet (Appendix VII-3).

GP addresses in England were obtained from the NHS Executive in 2001 after they carefully examined the research plan and objectives. Later an identical database including GP addresses was obtained from the National PCT Database, after the British Medical Association viewed the research summary plan and granted permission. The National PCT Database kindly negotiated access to this database. This second database was required since it included a further identifier code permitting linkage between general practices and the General Medical Services (GMS) Statistics databases.

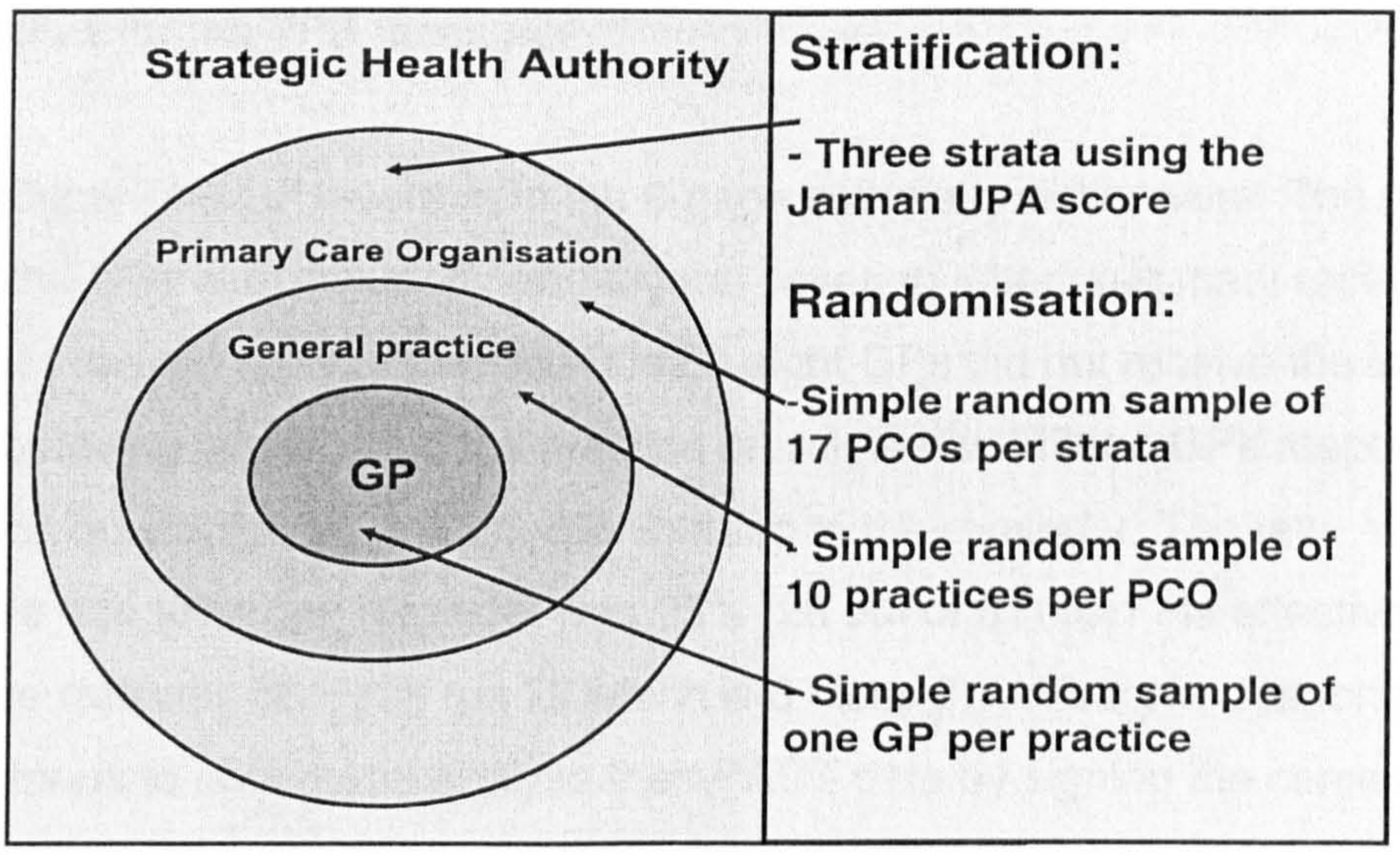
Participating GPs were asked to consent for us to access their statins and asthma prescribing data from PACT. Access to the PACT data for statins and asthma was negotiated with the Prescription Pricing Authority (PPA). Copies of signed consent forms were sent to the PPA. The original consent forms were not transferred since they were part of the confidential questionnaires. The PPA carefully assessed the consent forms and signatures and granted access to the PACT data for bronchodilators (BNF Sections 3.1), inhaled corticosteroids (BNF Section 3.2) and statins (BNF Section 2.12). PACT data were provided on paper (PACT 'catalogue'), comprising several pages of detailed prescribing information for each GP. Data were entered into computer

for analysis. Prescribing activity indicators were developed using the PACT data for the period April-June 2002.

7.2.2. Sampling and sample size

Sample size calculations indicated that 242 respondents were required for each survey (Chapter 6). Assuming 50% response rate and allowing for inaccuracies in the sampling frame, for each survey 510 GPs were sampled. The effective sample size was smaller owing to out-dated addresses or GPs on long term leave. It was known that about 2% of GP addresses in the GMS Statistics were inaccurate (Baker and Hann, 2001).

FIGURE 7.2. SAGE MULTI-STAGE SAMPLING SCHEME



The participants were identified using multi-stage stratified random sampling. Primary care trusts and primary care groups (PCOs) were stratified using Jarman scores. The Jarman scores were obtained from the National PCT Database, located at the University of Manchester (accessed October 2001). The PCOs were divided into three strata with equal number of PCOs; strata 1 (-36 < mean Jarman score < -9.8); strata 2 (-9.8 < mean Jarman score < +5);

and strata 3 (5 < mean Jarman score < 65). Then seventeen random PCOs were identified from each stratum. Ten GPs were randomly sampled from each PCO for each survey. If there were twenty or more practices in the PCO, only one GP was randomly sampled from each practice (Figure 7.2). Otherwise, more than one GP was randomly sampled from the required number of practices. In total ten practices (from five PCOs) received three questionnaires, 170 practices received two questionnaires (one asthma and one statins questionnaires) and the rest of the sampled practices (650) received only one questionnaire.

7.2.3. Regional pilot studies

Two separate 'regional' pilot surveys were conducted in Selby and York PCT to assess the sampling approach and response rate, and most importantly test the questionnaires. The results were also used to assess different analytical approaches for the TPB measures.

Eighty-one GPs were sent the 8-page asthma questionnaire. The packs sent to the GPs also included hand-signed covering letter, summary research plan and pre-paid reply envelope. At least eight GPs did not receive the letters owing to having moved from the practice or sick leave. Fifteen GPs responded to the first questionnaire and another thirteen to the reminder. The raw response rate after one reminder was 35% (28 out of 81) and the effective response rate was 38% (28 out 73 which is 81 less 8). All but two respondents agreed for us to obtain and analyse their PACT data by signing the consent forms.

A separate group of eighty-two GPs were sent the 8-page statins questionnaire. As with the asthma questionnaire, the packs also included signed covering letter, summary research plan and pre-paid reply envelope. At least five GPs did not receive the letters owing to having moved from the practice or sick leave. Sixteen GPs responded to the first questionnaire and another four responded to the reminder. The raw response rate after one

reminder was 24% (20 out of 82) and the effective response rate was 26% (20 out of 78). All but two respondents agreed for us to obtain and analyse their PACT data by signing the consent forms. There was no significant difference between the response rates of the two surveys (p=0.15).

7.2.4. Questionnaires

The questionnaires were developed following the guidelines provided (Conner and Sparks, 1996; Ajzen, 2002b). The process of questionnaire development is important part of conducting TPB surveys. In psychology this process is generally referred to as 'operationalisation' of the theory. 'Operationalisation' has been defined as the process of applying an abstract theory to a specific population, activity or setting (Resnicow et al, 2002).

Two separate questionnaires were developed for statins prescribing ('statins questionnaire', Appendix VII-1) and prescribing for the treatment of asthma ('asthma questionnaire', Appendix VII-2). As recommended (Conner and Sparks, 1996; Ajzen, 2002b), semi-structured interviews were used to elicit salient beliefs of GPs about outcomes, barriers to and facilitators of adhering to clinical guidelines for asthma drug treatment and for using statins (Chapter 4). An earlier version of the qualitative study's thematic framework (Table 4.1) was used for the identification of salient beliefs. All 'items' included in the framework were re-arranged under three main TPB themes of 'attitudes', 'subjective norms' and 'perceived controls'. Then specific analytical 'charts' were developed for each of the TPB themes. In each chart, the columns corresponded to the thematic items and the rows corresponded to the interviewees. The interviews were already transcribed and indexed for the main qualitative analysis. Indexed verbatim quotes from the transcribed interviews were copied and pasted into the charts. The charts were then analysed for identification of the salient beliefs for inclusion into the pilot questionnaires. Identification of salient beliefs followed a simple content analysis approach. Belief items were considered salient if more interviewees had reported the belief items, and the belief was perceived to be strong. Salient belief items were included in the pilot questionnaires.

Development of the final questionnaires then followed statistical 'item reduction' strategies using the results of the regional pilots (see 7.2.7). The TPB predictor measures (i.e. attitude, subjective norm and perceived control), behavioural intention and reported past behaviour were made compatible in terms of action, target, contexts and time, according to the 'principle of compatibility' (Ajzen, 1988; Conner and Sparks, 1996). Hence in all items the behaviour was defined as 'following clinical guidelines prescribing recommendation, in the practice over the following three months'. As recommended, the questionnaires incorporated multiple measures for each variable to increase measurement reliability (Bagozzi et al, 1989; Conner and Sparks, 1996). The preliminary questionnaires were assessed in terms of face validity and readability by a GP, a health psychologist with experience of TPB questionnaires and three health services researchers (Streiner and Norman, 2003). Then the questionnaires were formally piloted on two separate groups of GPs working in Selby and York PCT. The asthma and statins questionnaires were randomly assigned to the GPs, stratified by practice. The regional pilots' results were used to reduce the number of questions included in the final version of the questionnaire (7.3.1). Statistical techniques were used for item reduction (7.2.7).

The formats of the questionnaires used in the main surveys were as follows. Non-TPB items included two questions on self-identity, asking whether the GPs considered themselves as evidence-based, or patient-centred practitioners. Each questionnaire also had a series of demographic questions. The asthma and statins questionnaires included 46 and 47 TPB items respectively. TPB items were measured on 7-point bipolar or unipolar scales (Montano and Kasprzyk, 2002).

Dependent TPB variable

Behavioural intention. Three measures of intention were used in the questionnaires. Each item assessed one of the three main aspects of intentions: intention (or volition), desire and expectation (Conner and Sparks, 1996).

Independent TPB variables

Past behaviour. Two questions asked whether GPs' past prescribing (in the preceding three months) of statins and drugs for asthma were in line with clinical guidelines recommendations.

Attitude. This was measured both directly and indirectly. Six items were used for direct measurement of attitude using the semantic differential scaling approach (Oppenheim, 1996; pp 236-239). The items included the *good-bad* scale to capture general attitude, scales with 'instrumental' qualities (e.g. *valuable-worthless*) and scales with 'experimental' (e.g. *appropriate-inappropriate*) qualities (Ajzen, 2002b). To measure indirect attitudes, five salient perceived outcomes (beliefs) of adhering to secondary prevention of CHD and asthma clinical guidelines prescribing recommendations were identified from the qualitative interviews. For each outcome two questions were asked: the strengths of the GPs' belief that the outcome might occur as a result of prescribing ('belief strength'); and their evaluation of how good or bad it was if the outcome happened ('outcome evaluation'). The belief strength sub-items were assessed using *unlikely-likely* scales and outcome evaluations sub-items were assessed by *extremely bad-extremely good* scales. The responses to the two questions were multiplied and the mean of the sums of products of five multiplications constituted the final score for belief based (indirect) measure of attitude (Ajzen, 1991).

Subjective norm. Subjective norm captured perceived social norm or pressure for performing or not performing prescribing. Subjective norm was measured *directly* using two items with 'injunctive' quality. 'Injunctive' quality referred to the perceived view of the respondent that whether others expected or approved his or her behaviour (Ajzen, 2002b). In here the GPs' perceived view that whether important others expected or approved his or her prescribing in accordance with guideline recommendations was sought. *Indirect subjective norm* was also measured. It was based on normative beliefs. The beliefs were how 'important others' (salient referents) thought about prescribing according to the clinical guideline. The 'important others' (i.e. practice nurse, GP colleague, local hospital consultant and PCO prescribing adviser) were identified from the

interviews. Normative beliefs about 'important others' were assessed using two sets of sub-items. First set of sub-items measured the perceived views of the important others (e.g. practice nurse) about prescribing in accordance with guideline. These sub-items were called as normative '*belief power*'. They were assessed using *I should-I should not* scales. Second set of sub-items asked about the respondent *motivation to comply* with the perceived view of 'important others' (e.g. practice nurse). They were measured using '*not at all-very much*' scales. The 'belief power' and 'motivation to comply' sub-items were then multiplied to obtain the normative belief score corresponding with each 'important other'. For example if a GP perceived that his practice nurse colleague strongly expected him to prescribe statins for secondary prevention of CHD, and the GP also respected the view of the practice nurse colleague, he would have had positive normative belief. Similar questions were asked for all 'important others'. The mean of the normative beliefs' scores yielded the indirect subjective norm measure (Ajzen, 1991).

Perceived behavioural control (perceived control). Perceived control was intended to capture the respondents' confidence in performing the behaviour. As with previous independent variables it was measured by direct and indirect variables. Directly measured perceived control was captured using three items. The items were designed to be able to capture self-efficacy of GPs in performing prescribing (i.e. the level of perceived difficulty in adhering to the guideline prescribing recommendations) and whether they perceived they had control (Ajzen, 2002b) over prescribing (i.e. other factors did not prevent them in doing so). Indirect perceived control was measured based on six salient control beliefs. For each control belief the respondents were asked to respond to a 'belief strength' sub-item and a 'perceived power' sub-item. Different scales were used for the assessment of belief strength, but most used *strongly disagree-strongly agree* scales. For example they were asked how strongly they agreed that the BTS asthma guidelines' recommendations were based on evidence. Then the GPs were asked whether they perceived the control belief as a powerful source of influence on their behaviour ('perceived power'). For example they were asked to rate how easy it was to prescribe as guideline recommended if the guideline was evidence-based. These sub-items were assessed by scales such as *difficult-easy*. The score for indirect perceived

control was the mean of sum of products of control belief strengths and their perceived powers (Ajzen, 1991).

7.2.5. Prescribing outcome indicators

Source of prescribing data

PACT was the source of prescribing data used in the study. STAR-PU was considered as the most appropriate prescribing unit for these analyses as different gender specific age groups were weighted according to the specific therapeutic groups (i.e. BNF sections; Lloyd et al, 1995a). For example the weighting for males between 5 and 14 years for bronchodilators (BNF section 3.1) was 3.2, while for the same gender-age group the weighting for inhaled corticosteroids (BNF section 3.2) was 4.8 (Prescribing Support Unit, 2002). The PACT data used in this study focused on GPs rather than practices, because obtaining practice PACT data required the consent of the senior partners or all partners worked in the practices. On the other hand 'practice' demographic distribution was obtained from GMS Statistics. Therefore, STAR-PUs were calculated for the practices' list sizes, and then were divided by the number of whole time equivalent GPs (obtained from questionnaires) working in each practice. For each of asthma and statins surveys two primary prescribing outcomes were defined: 'effective delivery' and 'efficiency' indicators of prescribing. It was known that PACT based indicators were more accurate for efficiency measures than they were for quality measures (Majeed et al, 1997).

Defined Daily Dose, Adequate Daily Quantity and Net Ingredient Cost

Prescribing outcomes could be based on Defined Daily Doses (DDD; WHO Collaborating Centre for Drug Statistics Methodology, 1999). DDD is '*the assumed average maintenance dose per day for a drug used on its main indication in adults*'. DDD is not the recommended dose and in some cases may not be a real dose. It is rather a unit of measurement. DDD values as calculated by the WHO are based on the international experience. The advantage of using

DDD values rather than actual drugs prescribed is that it enables one to compare prescribing practices of similar drugs that are part of one Anatomical Therapeutic Chemical group (WHO Collaborating Centre for Drug Statistics Methodology, 1999). Since there are differences between the prescribing routines in the UK and the international values used for the calculation of DDDs, a separate set of values are calculated to reflect UK prescribing routines. There are also medicinal products in use in the UK for which there are no DDDs (e.g. topical drugs or some combination asthma inhaler drugs). Hence the Prescribing Support Unit funded by the Department of Health has defined Adequate Daily Quantities (ADQs) for use in the UK. Like DDD, ADQ is not a recommended dose but an '*analytical unit produced in order to compare more accurately the prescribing activity of primary care practitioners*'.

TABLE 7.1. DDD AND ADQ VALUES FOR STATINS AND INHALED CORTICOSTEROIDS

BNF (generic) name	Other names	Administration route	DDD	ADQ	Unit
Statins					
Atrovastatin	Lipitor	Oral	10	10	Mg
Cerivastatin (withdrawn)	Lipobay	Oral	200	100	Mcg
Fluvastatin	Lescol	Oral	40	30	Mg
Pravastatin	Lipostat	Oral	20	15	Mg
Simvastatin	Zocor	Oral	15	15	Mg
Inhaled corticosteroid					
Beclomethasone & salbutamol 50mcg/100mcg	Ventide	Inhaled aerosol		2	Puffs
Beclomethasone & salbutamol 100mcg/200mcg	Ventide	Inhaled powder		4	Rotacaps
Beclomethasone & salbutamol 200mcg/400mcg	Ventide	Inhaled powder		4	Rotacaps
Beclomethasone		Inhaled aerosol/powder	0.8	0.4	Mg
Beclomethasone CFC free		Inhaled aerosol		0.2	Mg
Budesonide		Inhaled suspension	1.5	1.5	Mg
Budesonide		Inhaled aerosol/powder	0.8	0.4	Mg
Budesonide & formeterol	Symbicort	Inhaled aerosol/powder		2	Puffs
Fluticasone propionate		Inhaled aerosol/powder	0.6	0.2	Mg
Fluticasone propionate		Inhaled suspension		2	Mg
Fluticasone & salmeterol	Seretide	Inhaled aerosol/powder		2	Puffs

Table extracted from Prescribing Support Unit, 2002, pp 31 and 35.

Both ADQ and DDD values were considered for the analysis of data in this study (Table 7.1). DDD values were used for statins. Statins DDDs represented equipotent doses of statins (Prescribing Support Unit, 2002). The DDD values for statins were obtained from the two latest PSU reports (Prescribing Support Unit, 2000; Prescribing Support Unit, 2002). The PSU updated some statins ADQs in a note published on 13 August 2004. The updated values were generally higher than the previous ones (e.g. ADQ for simvastatin increased from 15mg to 20mg). Statins ADQs were not used in this study since they did not reflect equipotent doses of statins but the current practice in the UK (Prescribing Support Unit, 2002). ADQs were used for inhaled corticosteroids. There were no DDDs available for some inhaled steroids. As inhaled corticosteroids' ADQs generally represented the equipotent doses of these drugs, they were useful for the analyses (Prescribing Support Unit, 2002).

Prescribing costs were calculated using net ingredient cost (NIC). NIC is cost of a medicine before any discounts or exemptions and it excludes any dispensing costs or fees. It also excludes any adjustment for income which may apply to prescription charges paid to the dispenser or to whether the patient has purchased a pre-payment certificate (National PCT Database, 2004).

Statins indicators

Statins DDDs per lipid lowering STAR-PU was suggested as 'effective delivery' indicator of prescribing in primary care (Department of Health, 2001b). The Department of Health (2001) also suggested three further effective delivery indicators of statins prescribing: 'number and percentage of patients discharged from hospital with a diagnosis of acute myocardial infarction prescribed statins', 'proportion of patients with a CHD risk greater than 30% over 10 years with a current prescription for treatment with statins' and 'proportion of patients with coronary artery disease prescribed a statin within the last 12 months'. None of these further indicators were obtainable from PACT data. For this study two statins prescribing outcomes were used. The effective delivery outcome was defined as statins DDDs per lipid lowering STAR-PU (Prescribing Support Unit,

2002) as weighted by the number of whole time equivalent GPs working in the practice. This outcome was simply called 'statins DDD per weighted STAR-PU'. The efficiency outcome was defined as statins cost per statins DDD and called as such.

Asthma indicators

Prescribing for asthma takes a variety of forms and preparations are prescribed in different quantities, administration forms and combinations. There are many generic and non-generic alternatives for the drugs. Different measures are used as effective delivery indicators for asthma prescribing. The Department of Health suggested '*proportion of patients with a diagnosis of asthma and a prescription for inhaled bronchodilators, who also received a prescription for inhaled corticosteroids within the last year*' as the indicator of effective delivery of asthma care (Department of Health, 2001b). A similar indicator was defined based on long acting β_2-agonists. Neither indicator could be obtained from PACT data.

Traditionally, a commonly used effective delivery indicator of asthma prescribing in the literature was the ratio of prophylactic prescribing (i.e. inhaled corticosteroids) to inhaled bronchodilators (Audit Commission, 1994; Majeed et al, 1997). The assumption was that bigger ratios reflected better prescribing. The Audit Commission (1994) suggested that '*those GPs who prescribe these preventive drugs less than one quarter as often as they prescribe bronchodilators*' needed to be scrutinised for their prescribing. This arbitrary threshold of 'a quarter' was a reflection of practice patterns at the time, rather than being based on evidence. Later studies suggested that the mean ratio as measured in DDD had increased from 0.49 to 0.73 (Majeed et al, 1999). The increase happened in about 6 years (1992-1998) and reflected faster increase in inhaled corticosteroid prescribing (Majeed et al, 1999; Premaratne et al, 1999). These ratios could be measured in terms of the ratio of prescribed DDDs, the ratio of NICs or the ratio of the number of dispensed items. Dispensed item was defined as one preparation on the prescription (Majeed et al, 1997). A prospective study of prescribing from 1993-1996 in West Midland assessed the validity of the ratio of DDDs (Frischer et al, 1999). They used

prescribing data from the General Practice Research Database (GPRD) and concluded that the ratio explained up to 37% of variation in asthma admission rates. It should be noted that data from the GPRD were linked to individual patients and their diagnoses. The prophylactic to bronchodilator ratio as measured from PACT data had been used successfully in previous studies in terms of NIC, number of prescribed items and DDDs (Feder et al, 1995; Griffiths et al, 1996; Griffiths et al, 1997; Aveyard, 1997). Use of number of dispensed items was not a valid measure of prescribing (Bogle and Harris, 1994) as it was dependant on the GPs' routines of managing their repeat prescription orders. For example GPs who renewed patients' repeat prescriptions every month prescribed twice the number of items as GPs who renewed repeat prescriptions bimonthly; even if both prescribed exactly the same medicine and of the same quantities. Also others had failed to find any relationship between quality of asthma care and the number of prescribed items (Jones et al, 1995).

The prophylactic to bronchodilator ratio was more valid if the ratio was measured in DDDs (Shelley et al, 1996; Majeed et al, 1997). However, one study concluded that the PACT based ratio of inhaled corticosteroids to bronchodilators in DDDs at health authority and general practice levels did not correspond with the admission rates for asthma (Shelley et al, 1996). Later the same group of researchers compared the asthma symptoms of patients from two general practices that differed in their corticosteroid to bronchodilator ratio. They concluded that the ratio was a valid indicator of quality (i.e. bigger ratio was indicative of less asthmatic symptoms in patients) if the difference was big enough (0.24 versus 1.53 in their study; Shelley et al, 2000).

The prophylactic to bronchodilator indicator however did not reflect recent developments in treatment of asthma and innovations in drug manufacturing and packaging. As noted before long-acting β_2-agonist bronchodilators were indicated for prophylactic treatment. Also more and more combination therapies of corticosteroids plus long- or short-acting β_2-agonists were manufactured and prescribed. All these combined preparations were categorised in BNF Section 3.2 with inhaled corticosteroids. It should also be noted that the STAR-PUs for inhaled bronchodilators differed with STAR-PUs for inhaled corticosteroids (Prescribing Support Unit, 2002). These might

explain why the ratio was not listed among the Department of Health recommended indicators. The Department of Health (2001) instead proposed using inhaled corticosteroids' DDDs per STAR-PU as indication of prescribing practice in accordance with BTS asthma guidelines' recommendations. It also proposed using inhaled corticosteroid NIC per DDD as 'efficiency' indicator of asthma care in primary care (Department of Health, 2001b). Likewise the National PCT Database used inhaled corticosteroids per corresponding STAR-PU as the effective delivery indicator of asthma treatment and cost per DDD indicator as the measure of efficiency. Simpler approaches had also been tried for the identification of better asthma services. One survey of general practices suggested that practices with better management of asthma were likely to have higher respiratory drug costs (including inhaled corticosteroid costs) as measured from PACT data (Jones et al, 1995).

In this study inhaled corticosteroid ADQs (Table 7.1) per weighted STAR-PU was used as the effective delivery measure of asthma prescribing. And inhaled corticosteroid cost per ADQ was used as the efficiency measure of asthma prescribing. Apart from the two primary prescribing outcome indicators, one further indicator of effective delivery was used as the secondary outcome (Griffiths et al, 1996; Griffiths et al, 1997): the ratio of inhaled corticosteroid costs (BNF Section 3.2) to inhaled β_2-agonists costs (BNF Section 3.1.1).

7.2.6. Non-response analysis

GMS Statistics were used for non-response analysis. GMS Statistics provides data about GPs, their practices and partnerships, services and patients. It is collected twice per year and is available through National PCT (formerly PCG/T) Database, National Primary Care Research and Development Centre, University of Manchester for research use. The GMS Statistics as presented in the National PCT Database are linked to the practices and not to the GPs. Although it was possible to link GPs with practices using an identifier code, it was not possible to link the sampled GPs with GMS Statistics directly. This was attributable to the surveys' sampling approach in which more than one GP was

sampled from some practices. The problem was overcome by separating those 'extra' GPs per practice from the rest and linking different files separately with the GMS Statistics data. Then the resulting files were merged back together and used for non-response analysis. T-tests and chi-squared tests were used for univariate non-response analysis. Multivariate analysis of non-response was performed using logistic regression. Linking and merging of files were performed using different computer software packages Microsoft Excel, Microsoft Access and SPSS at different stages. All the analyses were performed by the SPSS software package Version 10.

7.2.7. Analysis

Combining regional pilots and mains surveys

For the final analyses data from regional pilots and main surveys were combined. This was to increase analytical power and was conducted after careful examination of potential differences between the two samples. To achieve this, individual questions within the main surveys were matched with their corresponding questions in the regional pilots. Responses to the pilot questions were re-ordered if necessary to ensure they matched the main survey questions (e.g. if an item was scaled as 'definitely-definitely not' for the main survey and on the reverse scale for the pilot, the responses to the latter were re-ordered). Regional pilots' questions that had not been used in the main survey questionnaires were excluded from the final analyses.

Item ordering, missing values and scaling

All the items were ordered so that lower scores represented 'negative' responses towards the behaviour (e.g. 'definitely not' or 'strongly disagree'). Missing items were replaced by the mean of other items in the scale. In the first instance, attitude, subjective norm and perceived control items were treated bipolarly (-3 to 3), and intention and reported past behaviour items were treated unipolarly (1 to 7). The corresponding scales' scores were calculated by

summing the items' scores and dividing the total by the number of items in the scale. The indirect measures of attitude, subjective norm and perceived control were constructed as the means of sums of products of corresponding belief based items (Ajzen, 1991; see Table 7.2).

TABLE 7.2: CONSTRUCTION AND SCALING OF INDIRECT TPB MEASURES

TPB measure	Construct	Scaling approach
Indirect attitude (behavioural beliefs)	∑ (belief strength× outcome evaluation)	unipolar – bipolar
Indirect subjective norm (normative beliefs)	∑ (belief strength × motivation to comply)	bipolar – unipolar
Indirect perceived behavioural control (control beliefs)	∑ (belief strength × perceived power)	bipolar – unipolar

Item reduction

The results of the regional pilots were used to reduce the lengths of the questionnaires. Two analytical approaches were used for item reduction: the internal consistency approach or Likert scaling (Oppenheim, 1996, pp 195-200) and factor analysis (ibid, pp 166-171, 200). For the internal consistency approach Cronbach's α values for the scale were calculated. Then Cronbach's α was calculated for the scale if each item was deleted. If the deletion of an item from the scale resulted in an increase in the internal consistency of the scale (i.e. bigger Cronbach's α) then the item was a candidate for removal from the scale. Principal component analysis was used for identification of factors in the scales. Correlation matrices were inspected for the presence of coefficients of higher than 0.3. A scale considered factorable if the Kaiser-Meyer-Oklin value exceeded 0.6 and the Barlett's Test of Sphericity reached the statistical significance (Pallant, 2001). These statistical tests assessed the factorability of the scale. If the tests were successful, a minimum number of factors that explained majority of variance in the scale were kept in the analysis. Then the factors were subjected to rotation (Streiner and Norman, 2003). Rotation helped the items to load only on one factor; and also helped all the items on each factor to have the same sign (negative or positive). For directly measured components, attention was also paid to the endorsement rate in response to individual items (Streiner and Norman, 2003). If there was high endorsement in

response to an item (i.e. most responses were limited to one option), then the item was dropped from the final questionnaire or re-worded.

Other univariate analyses

Internal reliability of direct or indirect TPB variables (7.2.4) was assessed by measuring Cronbach's α (Bland and Altman, 1997). Pearson's correlation coefficients were calculated to examine univariate correlations between variables. T-tests were used for comparison of mean values between regional pilot and main surveys respondents, and between subgroups.

Regression analysis

Multivariate analyses were conducted using ordinary least square regression (Tabachnick and Fidell, 2001). These analyses were used to explain the variation in prescribing intentions and prescribing indicators (behaviour). F-tests in linear regression are robust and moderate departures from the analysis assumptions generally have little effects on the validity of null hypothesis tests (Cohen et al, 2003). Nonetheless following approaches were used to assess the analysis assumptions. Multi-collinearity was assessed by observing the variance inflation factor for each independent variable. Variance inflation factors of less than two with collinearity diagnostics of within recommended ranges (e.g. 'condition index' < 15) were representative of the absence of multi-collinearity. 'Casewise' diagnostics function was employed for the identification of outliers with residuals outside three standard deviations. Heteroscedasticity was assessed by observing the normality of the error term using histograms of standardised residuals, normal P-P (expected cumulative probability versus observed cumulative probability) plots and scatter plots of regression 'studentised deleted residuals' against standardised predicted values (Cohen et al, 2003). Cook's distance value was plotted against 'centred leverage value' for the identification of cases with strong influence on the regression model. Stepwise regression was not used (Miles and Shevlin, 2001).

Development of regression models

TPB regression models to explain variation in intention were developed as recommended by Ajzen (1991, 2002). For theoretical reasons all independent variables were included in the models at the same time. Non-significant variables were kept in the models unless they caused multi-collinearity.

Several variables were considered for inclusion in the regression models of prescribing indicators. These variables included TPB measures (i.e. attitude, subjective norm, perceived control and intention), TPB belief-based items and demographic variables. Apart from the TPB measures, there were no a priori strong reasoning to suggest which TBP belief items and demographic variables should be expected in the models. Also the sample size was small for reliable stepwise regression analysis in which statistical procedures could be used for the identification of important explanatory variables. Sample sizes in the magnitude of several hundred were required for such analyses. Therefore it was important to follow procedures to prevent 'over-fitting' models, avoid multi-collinearity and reduce the chances of losing important observations due to lack of power. Univariate analyses were performed to assess correlations between TPB belief measures and self-identity variables with prescribing indicators. Variables with univariate p-values of equal to or less than 0.2 were included in linear regression models to explain variation in prescribing outcomes. This arbitrary higher level of Type-I error was considered to reduce chances of important Type-II errors due to lack of power and theoretical understanding. At the next stage univariate analyses assessed relationships between practice and demographic variables and prescribing outcomes. Eligible variables were then included into the models following above criteria. Variables remained within the models if p-values for their corresponding regression coefficients were equal to or smaller than 0.2.

Transformation

Statistical transformation of data was avoided to prevent difficulty for the interpretation of the findings. Highly skewed variables or variables with similar values for many respondents (consultation length, practice deprivation score,

number of WTE GPs in practice) were transformed (square root or logarithm) or replaced with binary variables (based on median).

All the analyses were performed by the SPSS software package Version 10.

7.2.8. Ethical considerations

The study was conducted with careful consideration of ethical issues. Consent forms were included on the first page of the questionnaires and the participants were asked to tick relevant boxes and sign the form. The signature was required by the PPA to grant access to the GPs' prescribing data. Each questionnaire or reminder sent included hand-signed covering letter (example in Appendix VII-4), one page summary protocol and pre-paid return envelope. Questionnaires sent for the main surveys were also accompanied with 'questions and answers' page (Appendices VII-3). The returned questionnaires were appropriately coded to permit identification of the respondents (for PACT data and demographic information) and also to send reminder letters. The data from questionnaires was transferred to the computer. The questionnaires were stored securely. All the data received was dealt with confidentially. The participants were assured that the research reports would not contain any specific details which would allow identification of specific GPs or practices. The names and addresses of the GPs were obtained from the Primary Care Research Centre, University of Manchester after they sought the approval of the BMA on our behalf. For this purpose summary research plan was provided and it was understood that the addresses would not be used for other purposes. Local Research Ethics Committee approvals were not sought as at the time of data collection surveying health professionals did not require ethics approval. All respondents were told that they would be included in a lottery and the winner would receive a digital camera as a token of gratitude. After the surveys were accomplished the lottery was conducted in the presence of Prof Gillespie, Dean, Hull-York Medical School and as a result a digital camera was awarded to the winner. All the respondents were sent letters informing them of the result of the lottery.

7.3. Findings

7.3.1. Item reduction

The regional pilots' questionnaires were long and included several items for each TPB component, e.g. six separate items measured the respondents' intention. Internal reliability analysis and principal component analysis were applied to the regional pilots' results for item reduction. Despite small sample sizes behavioural beliefs and normative beliefs were factorable but control beliefs were not. The statistical analyses were complemented with interpretive judgements. This was to ensure the resulting questionnaire included important belief based items extracted from the interviews and covered different theoretical structures suggested by Ajzen (2002) for the directly measured components of the TPB (i.e. intention, attitude, subjective norm and perceived control). For example out of six intention items used in pilot questionnaire, three were kept for the main surveys. The retained intention items covered three theoretical aspects of intention: 'intention', 'desire' and 'expectation' (Conner and Sparks, 1996). The pilot respondents' comments on individual items were also taken into account.

As a result of item reduction the numbers of behavioural beliefs in asthma and statins questionnaires were reduced from seven in the regional pilots to five in the main surveys. Similarly the normative beliefs were reduced from five to four. There were ten control beliefs in asthma and eleven in statins pilot questionnaires. These were reduced to six control beliefs in the asthma main survey and to seven beliefs in the statins main survey. Item reduction also reduced the number of intention items from six to three, attitude items from twelve to six, subjective norm items from four to two and perceived control items from five to three in each questionnaire. In the regional pilot questionnaires there were five self-identity items (e.g. 'I see myself as an evidence-based practitioner'). Two self-identity items were kept in each of the main surveys (Appendices VII-1 and VII-2).

7.3.2. Optimal scaling of belief based items

The results of the regional pilot surveys were used to determine the optimal scaling approach for the belief based TPB items. The items used in the questionnaires were 7-point scales. The items were worded unipolarly (e.g. 'unlikely – likely', scored from 1 to 7) or bipolarly (e.g. 'bad – good', scored from -3 to +3). For the direct TPB measures (i.e. attitude, subjective norm and perceived control) it would make no difference in the statistical analysis whether the items were treated as bi-polar or otherwise, because numerical transformation of a unipolar scale to a bipolar scale just required subtracting by 4 which was a simple linear transformation.

Unlike direct variables, scaling of belief based variables could affect the results of statistical analyses. This was because the beliefs' composite scores were calculated by multiplying two related sub-items. For example each indirect attitude measure was calculated by multiplying the two components of the behavioural belief which were the belief strength and the outcome evaluation (Behavioural beliefs = $\sum$ 'behavioural belief strength' × 'outcome evaluation'). Because of this multiplication, the scaling approach affected the outcome. Suppose both belief strength and outcome evaluations were scaled bipolarly. If the respondent chose the lowest value for both items (i.e. -3), then the product was a large positive value (-3 × -3 = 9) suggesting the underlying belief had important positive contribution to the formation of attitude. If the items were treated unipolarly then the lowest value for each item was '1' and the product was also 1 (1 × 1 = 1), suggesting the underlying belief had minimal contribution to the attitude. And if one item was scaled unipolarly and the other item bipolarly, the product suggested negative contribution (1 × -3 = -3) in the attitude. These comparisons suggested that scaling of the items for indirect measures had important effects on the relationships between the TPB components. Unfortunately there was no a priori to decide the optimal scaling method for indirect measure items (Ajzen, 2002b). The choice of scaling approach should be based on the way the items were interpreted by the respondents (semantic scaling), which was in turn difficult to establish as different respondents interpreted the items differently. The other alternative was to test statistically different scaling approaches and choose the one that

provided the highest correlation values between the indirect and direct measures of the TPB components (Ajzen, 2002b). The results of the regional pilots were used to assess the optimal scaling method for the analysis of the main surveys. As the asthma regional pilot received higher response rate, primarily its results were used for establishing the optimal scaling approach unless the analyses were inconclusive.

Behavioural belief items

Conner and Spark (1996) suggested that for indirect attitude items the optimal scaling method was when both belief strength and outcome evaluation items were treated bipolarly. As mentioned before, Ajzen (2002) suggested there was no a priori method of establishing the optimal scaling method, but argued that for indirect attitude it was reasonable to treat belief strength as unipolar and the outcome evaluation as bipolar. His approach was in line with wording of the items in the survey questionnaires. The effects of different scaling approaches to the measurement of the indirect attitude component were assessed.

a. Unipolar – Unipolar scaling. All the behavioural belief items were constructed based on unipolar scaling of both belief strength and its outcome evaluation. The resulting items possessed acceptable internal reliability (Cronbach's α=0.85). The single variable of indirect attitude measurement was produced by adding the products. There was significant correlation of 0.55 between this variable and overall direct attitude variable.

b. Bipolar – Unipolar scaling. All the behavioural belief items were constructed based on bipolar scaling of belief strength items and unipolar scaling of outcome evaluation items (Cronbach's α=0.81). The products were added up to produce the indirect attitude measure. There was significant correlation of 0.44 between this variable and overall direct attitude variable.

c. Bipolar – Bipolar scaling (recommended by Conner and Spark, 1996). All the behavioural belief items were constructed based on bipolar treatment of belief strength and outcome evaluation items (Cronbach's α=0.78). The

correlation between the resulting indirect attitude measure and direct attitude measure was 0.47 and it was significant.

d. Unipolar – Bipolar scaling (recommended by Ajzen, 2002). The resulting scale produced the highest internal reliability measure (Cronbach's α=0.87). The indirect attitude measure also possessed the highest correlation coefficient with the direct measures of the attitude (r=0.63).

All the correlation values reported in options (a to d) above were with the direct attitude scale used in the regional pilots. The pilot attitude scale was constructed of 12 items. After item reduction procedures (see 7.3.1) only six of the items were kept for the final questionnaires. If the pilot's direct attitude scale was constructed by these six remaining items only, then scales according to b and c above were no longer significantly correlated with the direct attitude. The unipolar – bipolar construct ('d') retained a significant correlation coefficient of 0.57. These analyses supported unipolar scaling of belief strengths items and bipolar scaling of outcome evaluation item suggested by Ajzen (2002). This was also consistent with the wording of the items in the questionnaire.

Normative belief items

Conner and Spark (1996) suggested normative belief strength items should be scaled as bipolar and motivation to comply items as unipolar. This scaling approach fitted with the wording of the items used in the questionnaires and was followed for the analysis. The resulting items were internally consistent (Cronbach's α=0.77).

Control belief items

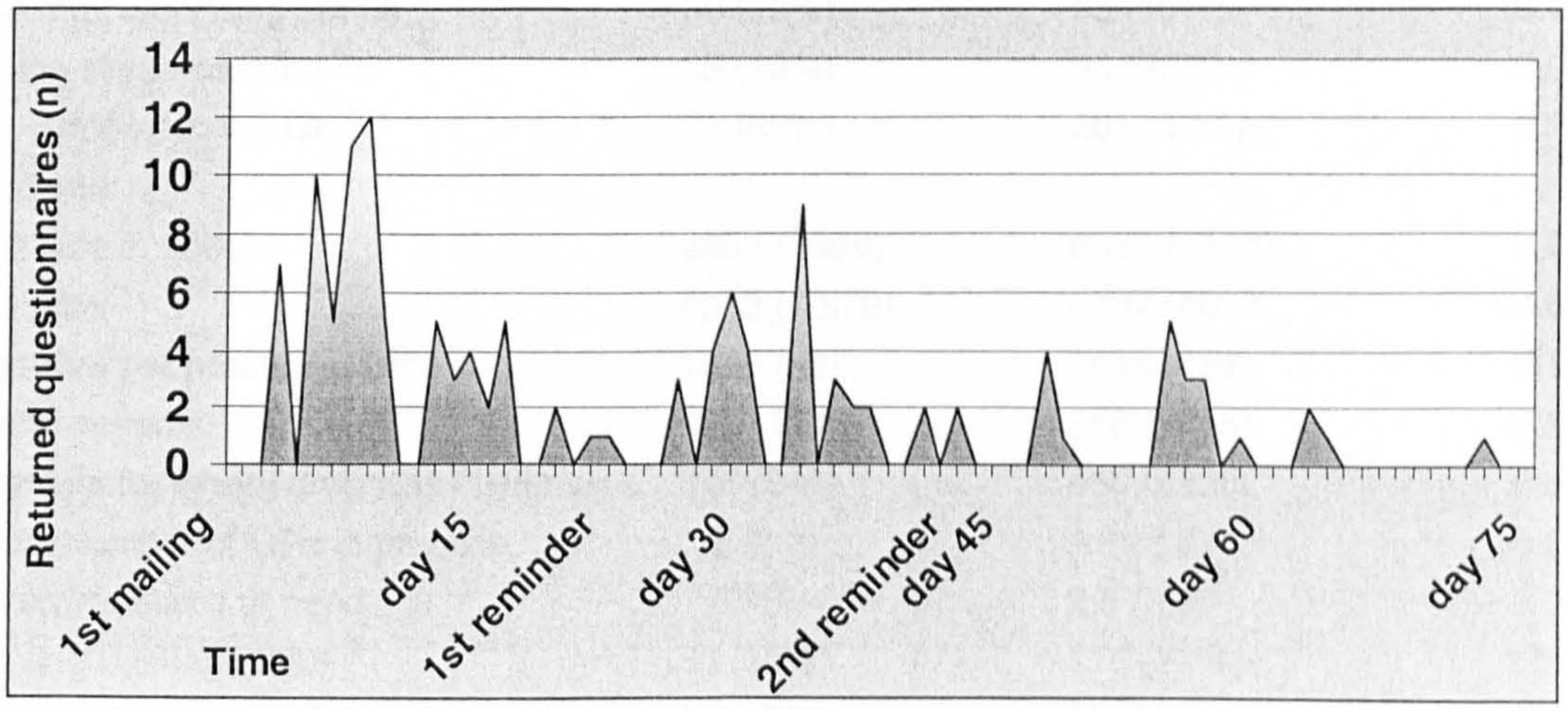
There were no recommendations on the best way of scaling the control belief items. Conner and Sparks (1996) suggested that future studies should focus on this. The results of the asthma and statins regional pilots were used to identify the optimal scaling approaches for control belief items (indirect perceived control). The results were inconclusive, perhaps because of the inherent difficulty of justifying the scaling approach for control belief items and the small

number of subjects responding to the pilots. Unlike behavioural belief items, it was not possible to conclude an optimal scaling approach for control belief items. Therefore, the results of the statins main survey were used to establish the optimal scaling approach for the analyses. The control belief strength items were scaled bipolarly for interpretive reasons. The perceived power items were scaled bipolarly and unipolarly in turn to identify the scaling approach that yielded the highest internal reliability. The reliability of a further approach was also assessed in which indirect perceived control measures were based on belief strength items only. Bipolar-unipolar scaling approach resulted in a scale with higher reliability (Cronbach's α=0.48) and higher correlation (r=0.35) with perceived control (direct measure).

7.3.3. Response rate, response pattern and non-response analysis

Despite efforts to boost response rate, the surveys achieved low response rates of 27% (135 out 500 which is 510 less 10) to the statins survey and 19% (94 out of 495 which is 510 less 15) to the asthma survey (Figure 7.3).

FIGURE 7.3. RESPONSE PATTERN TO THE STATINS SURVEY



The response rate to the asthma survey was significantly lower than the response rate to the statins survey (chi-squared=9.5; p=0.002). These response rates were lower than that achieved in the regional pilots with longer questionnaires, low quality prints and one reminder.

TABLE 7.3. UNIVARIATE ANALYSIS OF NON-RESPONSE

	Responders *	Non-responders *	P value
Two surveys combined			
List size in 2000	6673 (4219)	6409 (4112)	0.40
List size	6699 (4241)	6590 (4065)	0.73
List size per practicing GP	1877 (556)	1962 (585)	0.05
Rural patients	751 (1476)	659 (1375)	0.39
Patients for whom drug was dispensed	524 (1279)	409 (1171)	0.23
Total number of GPs in practice	3.8 (2.3)	3.5 (2.2)	0.17
Proxy measure of need	2.0 (3.2)	2.3 (3.4)	0.15
Having training status $	86 (38%)	198 (26%)	0.001
Having dispensing status $	48 (21%)	134 (18%)	0.27
Being single-handed $	43 (19%)	189 (25%)	0.05
All qualified in the UK $	148 (65%)	433 (58%)	0.06
Statins			
List size in 2000	6756 (4442)	6232 (4077)	0.21
List size	6827 (4483)	6423 (4028)	0.34
List size per practicing GP	1893 (538)	1981 (609)	0.14
Rural patients	822 (1669)	627 (1330)	0.23
Patients for whom drug was dispensed	627 (1464)	363 (1095)	0.06
Total number of GPs in practice	3.8 (2.3)	3.4 (2.2)	0.12
Proxy measure of need	2.1 (3.7)	2.3 (3.3)	0.44
Having training status $	53 (39%)	83 (24%)	0.001
Having dispensing status $	30 (22%)	55 (16%)	0.08
Being single-handed $	25 (19%)	95 (27%)	0.05
All qualified in the UK $	83 (62%)	201 (57%)	0.36
Asthma			
List size in 2000	6654 (3898)	6567 (4143)	0.98
List size	6513 (3878)	6737 (4097)	0.63
List size per practicing GP	1855 (584)	1945 (562)	0.17
Rural patients	649 (1145)	688 (1415)	0.81
Patients for whom drug was dispensed	376 (941)	450 (1235)	0.59
Total number of GPs in practice	3.7 (2.2)	3.6 (2.2)	0.66
Proxy measure of need	1.81 (2.4)	2.3 (3.6)	0.09
Having training status $	33 (35%)	115 (29%)	0.23
Having dispensing status $	18 (19%)	79 (20%)	0.93
Being single-handed $	18 (19%)	94 (24%)	0.37
All qualified in the UK $	65 (69%)	232 (58%)	0.05

* All values are means (SD) unless marked with $ where the values are numbers (%).

The responding GPs were not different from non-respondents in terms of the average number of patients registered with the practices (list size), practice proxy measure of need based on the Jarman score (Baker and Hann, 2001), total number of GPs working in the practices, number of rural patients registered with the practices, list size per practising GP in the practice, dispensing status and whether all the GPs in the practice were qualified in the UK (Table 7.3). The only clear difference observed at this stage was in the training status of the practices, which was predictive of higher response rate, especially to the statins survey. Marginally significant trends suggested GPs with larger list sizes, single-handed GPs and those practising among colleagues not qualified in the UK were less likely to respond to the surveys. Such findings might have been caused by multiple significance tests (Perneger, 1998). No adjustment was performed at this stage (Bland and Altman, 1995), but the findings were re-evaluated in multivariate analyses. The exclusion of GPs with practice list sizes of less than 500 (n=8) did not significantly change the results.

Multivariate analysis was performed using logistic regression. Combining the two surveys concluded that the training status of the practice (odds ratio=1.7; 95% CI=1.3-2.4; p<0.001) and the statins survey (odds ratio=1.7; 95% CI=1.2-2.3; p<0.001) were the only individual variables that contributed to higher response rates. For the statins survey, GPs within training practices were more likely to respond (odds ratio=2.1; 95% CI=1.4-3.2; p<0.001). For the asthma survey, GPs working in practices in which all the practitioners were qualified in the UK were more likely to respond (odds ratio=1.7; 95% CI=1.0-2.7; p=0.04).

Of those responding to the statins main survey 31% were female while the corresponding value for asthma survey was 44% (p=0.05). GMS Statistics provided the gender distribution of GPs in the practices, but no data about the gender of individual practitioners. Hence it was not possible to compare the gender profile of respondents with non-respondents. The difference was examined for single-handed GPs only (236 in total). There were no significant differences between the gender profile of single-handed GPs who responded to the questionnaires and those who did not.

The response rate from GPs working in practices in which all GPs were qualified in the UK (regardless of their ethnic origins) was higher than others. 59% of the sampled GPs were working in such practices, but 65% of the responding GPs were from them (p=0.06) (Table 7.3). These proportions were not evenly distributed between the two surveys as 69% of asthma respondents were from these practices (versus 58%; p=0.05), while the corresponding value for statins survey was 62% (versus 57%; p=0.36). This difference was partly due to the lower response rate of the GPs working in practices with non-UK qualified colleagues to the asthma survey. Further analyses were performed comparing GPs responding to the first mailing with the GPs that responded to first or second reminders. The two groups were compared for all the variables in Table 7.3. There were no differences between the two groups except for 'all qualified in the UK' variable (Table 7.4). The GPs who responded to the first mailing were more likely to be working in practices in which all the GPs were qualified in the UK (73%). In contrast only 55% of the GPs who responded to the two reminders were working in practices in which all the GPs were qualified in the UK (p=0.006). Within survey analyses revealed interesting phenomena: the difference between respondents to the first mailing and respondents to the reminders was not significant for the asthma survey (74% versus 64%; p=0.28), but the difference was significant for the statins survey (72% versus 49%; p=0.008).

TABLE 7.4. NON-RESPONSE ANALYSIS: RESPONSE TO THE FIRST MAILING AND REMINDERS AND ITS RELATIONSHIP WITH THE COUNTRY OF QUALIFICATION

	Asthma		Statins	
	First mailing	**Reminders**	**First mailing**	**Reminders**
All qualified in the UK	37 (74%)	28 (64%)	53 (72%)	30 (49%)
Not all qualified in the UK	13 (26%)	16 (36%)	21 (28%)	31 (51%)
Total	50 (100%)	44 (100%)	74 (100%)	61 (100%)
p value	0.28		0.008	

7.3.4. Demographic characteristics of GPs participating in the statins survey

In total 155 responded to the statins questionnaire, including twenty at the pilot stage of the study. The GPs had a mean age of 46 years. Years since graduation from medical school was calculated by subtracting the GPs' graduation year from 2001. On average they were 22 years from graduation. All but two GPs were principals. Eighteen (12%) GPs were single-handed; and 45 (29%) worked in practices with one or two GPs. Forty-four GPs were female, 39 GPs worked in dispensing practices and 52 GPs worked in practices with training status. Forty-five per cent of GPs reported that their practice formerly had fund-holding status. The GPs reported an average consultation time of 9 minutes.

TABLE 7.5. STATINS SURVEY PARTICIPANTS' DEMOGRAPHIC CHARACTERISTICS

	Total	Regional pilot	Main
Age (range)	46 (29-65)	44 (29-54)	47 (30-65)
Years since graduation (range)	22 (5-42)	18 (5-30)	23 (7-42)
Female (%)	44 (28)	2 (10)	42 (31)
Dispensing practice (%)	39 (25)	10 (53)	29 (22)
Training practice (%)	53 (34)	8 (40)	44 (33)
Former fund-holding status (%)	70 (45)	6 (30)	64 (48)
Computerised (%)	138 (89)	19 (95)	119 (89)
Consultation time (range)	9 (5-13)	9.5 (8-10)	9 (5-13)
Senior partner in practice (%)	71 (47)	14 (70)	57 (44)
WTE GPs working in practice	3.9	5.4	3.7

There were no significant differences between the GPs in the regional pilot and the main survey in terms of age and years since graduation from medical school. Consultation time for pilot GPs was on average (half a minute) longer than other GPs. The difference was of little importance, but it reached statistical significance. GPs in the regional pilot were more likely to be male ($p=0.05$), to work in dispensing ($p=0.004$) or larger ($p=0.003$) practices and to be senior partners in their practices ($p=0.03$; Table 7.5).

7.3.5. Intention to prescribe statins

Data from 155 questionnaires were analysed. The respondents expressed positive intentions to pursue guidelines' prescribing advice (three items; scale range: b1 to 7; mean=6.1; Cronbach's α=0.69). They also reported they had followed guidelines' advice for secondary prevention of CHD for more than half of their patients (two items; scale range: 1 to 7; mean=5.8; Cronbach's α=0.56). There were no significant differences between the regional pilot and the main survey respondents in terms of their intentions or reported prior prescribing.

TABLE 7.6. REGRESSION MODELS FOR EXPLAINING THE VARIATION IN STATINS' PRESCRIBING INTENTION

Dependent variable	β coefficient (CI)	p-value for β	Model's F-value (p-value)	Model's R squared
Model 1 (direct variables)			31.6 (<0.001)	0.41
Attitude	0.48 (0.29 to 0.67)	<0.001		
Subjective norm	0.13 (-0.01 to 0.27)	0.07		
Perceived control	0.15 (0.04 to 0.27)	0.01		
Model 2 (direct variables, past behaviour)			26.1 (<0.001)	0.44
Attitude	0.43 (0.24 to 0.62)	<0.001		
Subjective norm	0.12 (-0.02 to 0.26)	0.09		
Perceived control	0.12 (0.00 to 0.23)	0.05		
Past behaviour	0.20 (0.04 to 0.36)	0.02		
Model 3 (indirect variables)			25.0 (<0.001)	0.35
Indirect attitude	0.10 (0.05 to 0.13)	<0.001		
Indirect subjective norm	0.02 (-0.02 to 0.05)	0.35		
Indirect perceived control	0.06 (0.01 to 0.11)	0.02		
Model 4 (all TPB variables, past behaviour)			31.1 (<0.001)	0.48
Indirect attitude	0.07 (0.03 to 0.11)	<0.001		
Attitude	0.25 (0.05 to 0.45)	0.02		
Indirect perceived control	0.04 (0.002 to 0.08)	0.04		
Past behaviour	0.32 (0.17 to 0.48)	<0.001		

GPs reported very positive attitudes towards following statins prescribing recommendations for secondary prevention of CHD (six items; scale range: -3 to 3; mean=2.1; Cronbach's α=0.87) and perceived positive social norms (two

items; scale range: -3 to 3; mean=1.9; Cronbach's α=0.89). Their perceived ability to overcome barriers was lower (three items; scale range: -3 to 3; mean=1.1; Cronbach's α=0.58). Among the perceived control items, the mean response to the item assessing the 'controllability' of prescribing (whether GPs perceived it was under their control and not dependent on other circumstances) was 0.6. On the other hand they perceived the behaviour as less difficult in terms of decision making and performing ('self-efficacy'; mean=1.4).

Direct measures of attitude, subjective norm and perceived control explained up to 41% of the variance in prescribing intentions. The addition of prior reported behaviour increased this to 44% (Models 1 and 2 in Table 7.6). Among TPB measures, attitude was the most powerful explanatory variable for intention followed by perceived control. Self-identity measures did not improve the regression model.

Attitude, subjective norm and perceived control were also indirectly measured using belief based items (Table 7.7). All indirectly measured variables were positively correlated with the direct measures and the correlations coefficients varied from moderate (0.35) to large (0.65) (Cohen, 1988). The indirect TPB measures were included in separate regression analyses to explain the variation in intention. Indirect measures of attitude, subjective norm and perceived control explained 35% of the variation in intention. Like direct measures, the regression coefficients for indirect measures of attitude and perceived control reached statistical significant while the coefficient for indirect subjective norm did not. The addition of reported past behaviour to Model 3 increased the explained portion of variation in intention to 44% ($F=26.3$, $p<0.001$). Then a separate model was devised including all direct and indirect TPB variables. Variables with large p-values that were suspected of causing multi-collinearity were excluded. The resulting model explained 41% of variation in intention ($F=31.9$, $p<0.001$). Then the reported past behaviour was added to the model (Model 4; R-squared=0.48). None of subjective norm variables contributed to this final model (Table 7.6). Individual normative belief items were also assessed for their potential contribution to Model 4, but failed to improve. Self-identity measures did not improve the models. No other demographic or practice variable contributed to the model.

TABLE 7.7. INDIRECT BELIEF BASED MEASURES OF THE TPB, AND THE MODAL BELIEF ITEMS USED IN THE STATINS SURVEY AND THE MEAN RESPONSES TO THEM

	Correlation with corresponding direct measure: Pearson r (p value)	Internal reliability: Cronbach α	Mean (SD) (scales' range: -21 to 21)
Indirect attitude	0.52 (<0.001)	0.72	12.4 (4.7)
Indirect subjective norm	0.65 (<0.001)	0.80	10.7 (5.4)
Indirect perceived control	0.35 (<0.001)	0.48	7.4 (3.8)
Belief items			
Behavioural beliefs (range: 1 to 7)			
If I (GP) prescribe statins for secondary prevention of CHD as recommended in clinical guidelines ...			
1. Patients will be healthier			5.5 (1.5)
2. I receive 'quality markers'			4.7 (1.8)
3. Standard of care will not be judged negligent			4.8 (1.7)
4. It prevents harm to patients			5.2 (1.6)
5. It provides better quality of care for patients			5.6 (1.5)
Normative beliefs (range: -3 to 3)			
... thinks I should/should not prescribe statins for ...			
1. Practice nurse			1.9 (1.2)
2. GP colleague			1.9 (1.0)
3. Local hospital consultant			2.3 (1.0)
4. PCO prescribing adviser			2.2 (1.1)
Control beliefs (range: -3 to 3)			
1. Clinical guidelines for secondary prevention of CHD are evidence-based (facilitator)			1.9 (1.2)
2. Statins prescribing recommendations in guidelines have substantially changed			0.5 (1.5)
3. I am under time pressure to care for CHD patients			1.3 (1.9)
4. The inclusion guidelines for secondary prevention of CHD in the BNF is appropriate (facilitator)			2.3 (1.0)
5. NSF for heart disease advocates statins prescribing for secondary prevention of CHD (facilitator)			2.1 (1.3)
6. Prescribing statins as recommended could exhaust practice budget			1.5 (1.7)
Self Identity			
1. I am an evidence-based practitioner			5.2 (1.5)
2. I am a patient-centred practitioner			6.1 (1.1)

7.3.6. Statins prescribing

Primary prevention

GPs were asked to report their practice of statins prescribing for primary prevention of CHD. Only three GPs (2%) mentioned they did not prescribe statins for primary prevention. Twenty-eight GPs (18%) had no explicit criteria, 14 GPs (9%) prescribed for those with a minimum CHD risk of 15% over the next ten years, 44 GPs (29%) prescribed for patients with a CHD risk of 15-30% and 65 GPs (42%) prescribed for patients with a CHD risk of more than 30% over the next ten years. There was no identifiable pattern of difference between these groups in terms of their prescribing as measured by statins DDD per weighted STAR-PU (ANOVA test, between group degrees of freedom=3, F=1.3, p=0.28).

Effective delivery indicator: statins DDD per weighted STAR-PU

Data for this outcome was available for 128 GPs. Statins DDD per weighted STAR-PU ranged from 0.09 to 1.39 (median=0.61; mean=0.66; SD=0.31). Higher values were assumed to correspond with more effective CHD prevention practice (Department of Health, 2001b). There was no significant difference between the regional pilot and the main survey respondents (t=1.1; mean difference=0.10; p=0.27).

Model 1 in Table 7.8 reported the results of regression analysis using TPB measures as predictors. None of the direct or indirect TBP variables significantly contributed to the regression model. Following the criteria explained in 7.2.7, self-identity and TPB belief based variables were included in the analysis. The variables explained 14% of variance in the outcome. Regression coefficients for only one behavioural belief and one self-identity variable were significant. This suggested those who viewed themselves as evidence-based practitioners and those perceived that following guidelines recommendations for statins prescribing resulted in 'quality markers' had prescribed more statins

DDD per weighted STAR-PU. No other TPB variable met the pre-defined criteria to be included in the model.

TABLE 7.8. REGRESSION MODELS FOR EXPLAINING VARIATION IN STATINS PRESCRIBING – EFFECTIVE DELIVERY INDICATOR

Dependent variable	B coefficient (CI)	p for β	Model's F (p)	Model's R squared
Model 1 (TPB variables)			3.4 (0.007)	0.14
Behavioural belief 2: quality markers	0.008 (0.001 to 0.016)	0.028		
Self identity 1: evidence-based practitioner	0.042 (0.002 to 0.081)	0.040		
Behavioural belief 5: quality of patient care	-0.009 (-0.018 to 0.000)	0.056		
Control belief 4: inclusion in the BNF	-0.007 (-0.015 to 0.000)	0.064		
Normative belief 4: PCO adviser	-0.006 (-0.014 to 0.003)	0.18		
Model 2			11.1 (<0.001)	0.39
Practice deprivation score	0.23 (0.14 to 0.32)	<0.001		
Female GP	-0.20 (-0.31 to -0.097)	<0.001		
Years since graduation	0.008 (0.002 to 0.014)	0.008		
Control belief 4: inclusion in the BNF	-0.006(-0.013 to 0.000)	0.066		
Behavioural belief 2: quality markers	0.005 (0.000 to 0.011)	0.068		
Interest in CHD	0.06 (-0.03 to 0.16)	0.19		
Variables excluded from Model 2 because univariate p > 0.2				
Training status	Asthma clinic		CHD clinic	
Computerised practice	Previously fund-holding			
Variables excluded from Model 2 after initial inclusion (because p for regression coefficient > 0.2)				
Consultation length	Dispensing status		Senior partner	
All partners graduated in the UK	Number of WTE GPs in practice			

Then a second regression model was devised to assess the variance in effective delivery indicator. All variables from Model 1 (Table 7.8) plus demographic and practice variables were eligible for inclusion if they had a univariate p-value of smaller than 0.2 against the indicator and p-value remained less than 0.2 in multi-variate analysis (Table 7.8). Consultation time and practice deprivation scores were transformed to binary variable as they were skewed. The number of WTE GPs in practice was analysed as both binary

and continuous variables. 'Years since graduation' was used in preference to age. Model 2 of Table 7.8 represented the results of the final regression model after exclusion of ineligible variables. The model explained 39% of variation in the prescribing indicator. Given that there were no individual patient characteristics in the model, this was acceptable. The model suggested that GPs working in practices with higher deprivation (binary variable based on median) and GPs with more years in practice had prescribed more statins per prescribing unit. On the other hand female GPs had prescribed less statins as measured by the indicator. Sensitivity analysis was performed by testing the model on the respondents to the main survey only (excluding regional pilot). The results were similar to those reported in Table 7.8.

Efficiency indicator: statins cost per DDD

This indicator was meant to reflect cost conscious prescribing by GPs, in which GPs chose statins that were cheaper per statins DDD. Costs were for net ingredients measured in pounds sterling. Higher values for this outcome reflected less efficient prescribing. The data were normally distributed (range: 0.40 to 0.83, median=0.62, mean=0.62, SD=0.08). Three cases were excluded from the analysis as outliers. There was no significant difference between the regional pilots and the main survey respondents for this outcome (t=1.1, mean difference=0.02, p=0.28).

In the first regression model, the TPB variables with correlations smaller than 0.2 were included (Model 1). Indirect subjective norm and indirect control belief explained 4% of variance in efficiency indicator. This suggested that the GPs feeling under more social pressure from their 'important others' were likely to prescribe less efficiently. In reverse those perceiving higher levels of self-efficacy and control over barriers were spending less per statins DDD. The observed effects from both beliefs were small (Table 7.9). Then the control and normative belief variables with correlation coefficient p-values of less than 0.2 were added to the model (four belief based variables).

In presence of belief items, the indirect variables were no longer significant contributors and were excluded because of large p-values (Model 2).

Model 2 explained 12% of variance in the outcome. It suggested GPs perceiving pressure from PCO advisers for prescribing statins were less likely to prescribe efficiently. On the other hand GPs believing more strongly in NSF support for statins prescribing for secondary prevention were likely to prescribe cheaper statins (Table 7.9).

TABLE 7.9. REGRESSION MODELS FOR EXPLAINING VARIATION IN STATINS PRESCRIBING – EFFICIENCY INDICATOR

Dependent variable	β coefficient (CI)	p for β	Model's F (p)	Model's R squared
Model 1 (TPB variables)			2.7 (0.07)	0.04
Indirect perceived control	-0.004 (-0.008 to 0.000)	0.048		
Indirect subjective norm	0.003 (0.000 to 0.005)	0.048		
Model 2 (TPB belief based variables)			5.3 (0.002)	0.12
Normative belief 4: PCO adviser	0.003 (0.001 to 0.005)	0.003		
Control belief 5: NSF advocates statins prescribing	-0.002 (-0.004to -0.001)	0.006		
Control belief 3: time pressure	-0.002 (-0.005 to 0.001)	0.11		
Model 3 (all variables)			6.4 (<0.001)	0.29
Normative belief 4: PCO adviser	0.003 (0.002 to 0.005)	<0.001		
Senior partner	0.036 (0.011 to 0.060)	0.004		
Control belief 5: NSF advocates statins prescribing	-0.002 (-0.004 to -0.001)	0.008		
Number of WTE GPs in practice	0.007 (0.002 to 0.012)	0.010		
Interest in CHD	-0.028 (-0.052to -0.003)	0.027		
CHD clinic	-0.028 (-0.050 to 0.000)	0.048		
Control belief 3: time pressure	-0.003 (-0.006 to 0.000)	0.051		

Variables excluded from Model 3 because univariate p > 0.2

Number of WTE GPs in practice	Asthma clinic	Age
Computerised practice	Previously fund-holding	Practice deprivation score
All partners graduated in the UK	Gender	Dispensing status
Consultation length		

Variables excluded from Model 3 after initial inclusion (because p for regression coefficient > 0.2)

Practice list size	Training status	Years since graduation

Model 3 was developed by addition of demographic and practice characteristics to Model 2. Unlike the results observed for the effective delivery outcome, the TPB measures retained their contribution to the outcome in the presence of demographic and practice characteristics. Model 3 explained 29% of variance in the efficiency outcome. Apart from belief based items (all three significant), working in a larger practice (more GP colleagues) and being the senior partner in practice were both linked to more costly statins prescribing. In contrast, GPs who mentioned they provided CHD clinics in their practices and those with interest in CHD were more efficient prescribers (Table 7.9).

7.3.7. Demographic characteristics of GPs participating in the asthma survey

Total number of respondents to the asthma survey was 122 of whom 28 GPs participated in the regional pilot. All but one GP were principals in their practices. All but ten GPs were working in computerised practices, with fifty per cent reporting asthma clinical guidelines on their computer systems. Other practice characteristics were reported in Table 7.10.

TABLE 7.10. ASTHMA SURVEY PARTICIPANTS' DEMOGRAPHIC CHARACTERISTICS

	Total	Regional pilot	Main
Age (range)	46 (29-69)	45 (35-66)	46 (29-69)
Years since graduation (range)	21 (6-43)	22 (11-43)	21 (6-41)
Female (%)	51 (42)	11 (41)	40 (44)
Dispensing practice (%)	33 (27)	16 (59)	17 (19)
Training practice (%)	36 (30)	9 (33)	27 (30)
Former fund-holding status (%)	45 (37)	5 (19)	40 (45)
Computerised (%)	107 (88)	27 (100)	80 (89)
Consultation time (range)	9.6 (5-15)	9.8 (7-10)	9.5 (5-15)
Senior partner in practice (%)	55 (45)	18 (67)	37 (43)

The respondents' mean age was 46. On average, 21 years had passed since their graduation from medical school. Forty-two per cent of the respondents were female. Forty-five per cent of GPs reported that their practice formerly had fund-holding status. There were no significant differences between the GPs in pilot and main surveys in terms of age, gender, years since

graduation from medical school and consultation time. GPs in the pilot were more likely to be working in dispensing practices ($p<0.001$), in formerly fund-holding practices ($p=0.01$) and to be senior partners ($p=0.03$) in their practices (Table 7.10).

7.3.8. Intention to prescribe medicines for asthma

In total 122 GPs responded to the asthma questionnaire. GPs intended to follow clinical guidelines' prescribing recommendations for the treatment of asthma (three items; scale range: 1 to 7; mean=5.5; Cronbach's $\alpha=0.58$). The exclusion of the third intention item from the scale increased internal reliability coefficient to 0.79, but did not improve the regression model or the significance of regression coefficients. Hence the original scale of three items was used for the analyses. GPs also reported that they had followed guidelines advice when prescribing for asthmatic patients within the past three months for more than half of their patients (two items; scale range: 1 to 7; mean=5.4; Cronbach's $\alpha=0.54$). There were no significant differences between the regional pilot and the main survey respondents in terms of their intentions or reported prior prescribing.

Three direct TPB measures assessed GPs' perceived controls, attitudes and subjective norms. GPs perceived themselves to be able to overcome barriers to evidence-based prescribing for asthma (three items; scale range: -3 to 3; mean=1.5; Cronbach's $\alpha=0.78$). GPs were equally confident of their self-efficacy and control over factors that hindered this. Less than 10% of the respondents did not perceive themselves to have control over barriers (mean response of less than 0). GPs had positive attitudes towards following clinical guidelines' prescribing recommendations (six items; scale range: -3 to 3; mean=2.1; Cronbach's $\alpha=0.90$) and perceived that the social norm was to prescribe in that manner (two items; scale range: -3 to 3; mean=2.0; Cronbach's $\alpha=0.63$).

Up to 43% of variance in prescribing intentions was explained by direct TPB measures. Perceived control was the sole predictor variable with significant regression coefficient. The addition of prior reported behaviour did not significantly improve the model as the regression coefficient for prior behaviour was non-significant (Models 1 & 2 in Table 7.11). Self-identity measures did not improve the models.

TABLE 7.11. REGRESSION MODELS FOR EXPLAINING THE VARIATION IN INTENTIONS TO PRESCRIBE FOR ASTHMA TREATMENT AS RECOMMENDED BY CLINICAL GUIDELINES

Dependent variable	β coefficient (CI)	p value for β	Model's F value (p value)	Model's R square
Model 1 (direct variables)			26.5 (<0.001)	0.43
Perceived control	0.44 (0.26 to 0.62)	<0.001		
Attitude	0.20 (-0.03 to 0.42)	0.09		
Subjective norm	0.19 (-0.05 to 0.43)	0.11		
Model 2 (direct variables, past behaviour)			20.9 (<0.001)	0.44
Perceived control	0.40 (0.22 to 0.58)	<0.001		
Subjective norm	0.18 (-0.06 to 0.41)	0.15		
Attitude	0.13 (-0.10 to 0.37)	0.27		
Past behaviour	0.15 (-0.03 to 0.32)	0.10		
Model 3 (indirect variables, past behaviour))			34.9 (<0.001)	0.40
Indirect subjective norm	0.10 (0.07 to 0.13)	<0.001		
Past behaviour	0.21 (0.04 to 0.37)	0.02		
Model 4 (all TPB measures)			46.1 (<0.001)	0.47
Indirect subjective norm	0.07 (0.04 to 0.11)	<0.001		
Perceived control	0.44 (0.25 to 0.62)	<0.001		

Attitude, subjective norm and perceived control were also indirectly measured using modal belief based items (Table 7.12). Indirectly measured attitude and subjective norm were positively correlated with their direct measures (r=0.45 and 0.76 respectively). Six modal control beliefs were used for the assessment of indirect perceived control. Three belief items (control belief items 2, 3 and 5 in Table 7.12) were not correlated with the direct measure, while others had statistically significant correlations (r from 0.22 to 0.44). The inclusion of items 2, 3 and 5 in the indirect perceived control belief

resulted in slight reduction in the reliability of the scale (Table 7.12). There was statistically significant correlation (r=0.29) between two control belief items of 'evidence base' and 'hearing about the guideline' (control belief items 1 and 6).

TABLE 7.12. INDIRECT BELIEF BASED MEASURES OF THE TPB, AND THE MODAL BELIEF ITEMS USED IN THE ASTHMA SURVEY AND THE MEAN RESPONSES TO THEM

	Correlation with corresponding direct measure: Pearson r (p value)	Internal reliability: Cronbach α	Mean (SD) (scales' range: -21 to 21)
Indirect attitude	0.45 (<0.001)	0.73	12.4 (4.1)
Indirect subjective norm	0.76 (<0.001)	0.88	10.1 (6.1)
Indirect perceived control	0.33 (<0.001)	0.32	4.4 (3.6)
Indirect perceived control (3-beliefs) *	0.42 (<0.001)	0.39	4.6 (6.0)
Belief items			
Behavioural beliefs (range: 1 to 7)			
If I (GP) prescribe for treatment of asthma as recommended in clinical guidelines ...			
1. Patients will be healthier			5.5 (1.3)
2. I receive 'quality markers'			3.7 (1.8)
3. Standard of care will not be judged negligent			4.1 (1.7)
4. It prevents harm to patients			5.3 (1.5)
5. It provides better quality of care for patients			5.7 (1.3)
Normative beliefs (range: -3 to 3)			
... thinks I should/should not prescribe for asthma as recommended ...			
1. Practice nurse			2.0 (1.0)
2. GP colleague			1.7 (1.1)
3. Local hospital consultant			2.0 (1.2)
4. PCO prescribing adviser			2.1 (1.1)
Control beliefs (range: -3 to 3)			
1. BTS guidelines for asthma are evidence-based (facilitator) *			1.4 (1.5)
2. BTS asthma guidelines have substantially changed their recommendations			0.2 (1.4)
3. I am under time pressure to care for asthmatic patients			1.2 (1.8)
4. BTS asthma guidelines are not flexible *			0.6 (1.6)
5. The inclusion BTS guidelines for asthma in the BNF is appropriate (facilitator)			2.3 (1.2)
6. I constantly hear about the BTS asthma guidelines (facilitator) *			0.3 (1.6)
Self Identity			
1. I am an evidence-based practitioner			5.1 (1.4)
2. I am a patient-centred practitioner			5.7 (1.4)

*** Control beliefs that contributed to the 'indirect perceived control (3-beliefs)' variable.**

The indirect measures were included in separate regression analyses to explain variation in the TPB. Indirect perceived control had no significant contribution to variation in intention. This remained unchanged even when the measure was based on three beliefs 2, 3, and 5 discussed above. Indirect measure of attitude too had no contribution to the model. Unlike what reported for statins surveys, indirect subjective norms were the only significant contributors to the model. The analysis suggested significant multi-collinearity between indirect subjective norm and indirect attitude. The exclusion of indirect attitude slightly reduced R-squared and changed no other parameter. A model based on indirect subjective norm and reported past behaviour accounted for 40% of variation in intention (Model 3, Table 7.11). All TPB variables were eligible for inclusion in Model 4, but only two significantly contributed to it. The model including direct perceived control and indirect subjective norm explained 47%; and inclusion of past behaviour in the model did not significantly change the parameters (Model 4, Table 7.11).

Given that indirect measures of perceived control and attitude had no contribution to the models, individual belief items were included to identify control and behavioural beliefs that were linked to intention. One control belief item (control belief 2) was the only one that significantly improved Model 4 (R-squared change=0.03, F change=6.1, p=0.02). It suggested that those who agreed more strongly with 'BTS guidelines had changed their recommendations' had weaker intentions to follow those recommendations (regression coefficient=-0.031, p=0.015). Self-identity measures did not improve the models.

7.3.9. Prescribing for asthma

Effective delivery indicator: inhaled corticosteroid ADQ per weighted STAR-PU

Data for 96 GPs was available for the calculation of this indicator. Inhaled corticosteroid ADQ per STAR-PU ranged from 0.13 to 2.43 (median=0.91, mean=0.82, SD=0.49). GPs with higher values for this indicator assumed to

have followed the BTS asthma guidelines' recommendations more closely (Department of Health, 2001b). There was no significant difference between the regional pilot and main survey respondents for this indicator (mean difference=0.23, t=1.7, p=0.09).

TABLE 7.13. REGRESSION MODELS FOR EXPLAINING THE VARIATION IN EFFECTIVE DELIVERY INDICATORS OF PRESCRIBING FOR ASTHMA

Dependent variable	β coefficient (CI)	p for β	Model's F (p)	Model's R squared
A) Inhaled corticosteroid DDD per weighted STAR-PU				
Model 1 (all variables)			4.4 (0.003)	0.18
Computerised practice	0.44 (0.09 to 0.80)	0.02		
Years since graduation	0.01 (0.002 to 0.03)	0.02		
Asthma clinic	-0.23 (-0.43 to -0.03)	0.03		
Dispensing status	0.17 (-0.04 to 0.38)	0.11		
Variables excluded from Model 1 because univariate p > 0.2				
Senior partner	Training status		Consultation length	
Number of WTE GPs working in practice	Practice deprivation score		Age	
All partners graduated in the UK	Practice list size		Interest in asthma	
Variables excluded from Model 1 after initial inclusion (because p for regression coefficient > 0.2)				
Previous fund-holding status	Gender			
B) Ratio of prophylactic cost to bronchodilator cost				
Model 2 (all variables)			8.3 (0.005)	0.09
Control belief 2: BTS asthma guidelines have substantially changed their recommendations	-0.03	0.005		
Variables excluded from Model 2 because univariate p > 0.2				
Asthma clinic	Previous fund-holding status		Dispensing status	
Training status	Consultation length		Gender	
Senior partner	Practice list size		Age	
Years since graduation	Number of WTE GPs in practice			
Variables excluded from Model 2 after initial inclusion (because p for regression coefficient > 0.2)				
Practice deprivation score	Computerised practice		Interest in asthma	
All partners graduated in the UK				

No direct or indirect TPB variable significantly contributed to the regression model. Following the criteria explained in 7.2.7, only one behavioural belief and one control belief measure met the statistical criteria for inclusion in

the regression models. However, these three variables had no significant contribution to variation in the outcome indicator. The indicator was transformed to its square root yielding a distribution close to normal (range: 0.36 to 1.56; median=0.91; mean=0.92, SD=0.25). Procedures followed for the non-transformed variable were repeated. No significant univariate or multivariate relationships between the TPB measures and the square root of effective delivery indicator were observed. Therefore Model 1 was devised using demographic and practice characteristics only (Table 7.13). The included variables explained 20% of variation in inhaled corticosteroid DDD per weighted STAR-PU. The model suggested that more experienced GPs and those who worked in computerised practices were more likely to offer better quality of asthma prescribing. On the other hand the presence of asthma clinics was correlated with less inhaled corticosteroid DDDs per weighted prescribing unit.

As the analyses identified no relationship between the TPB measures and the primary effective delivery outcome, they were repeated using the secondary outcome. The prophylactic cost to bronchodilator cost ratio was the planned secondary effective delivery outcome (range: 0.6 to 4.2; median=1.8; mean=1.8; SD=0.67). It was assumed that GPs with bigger ratios had better quality of asthma prescribing. Three TBP variables met the criteria for inclusion into the regression model: indirect attitude, behavioural belief 5 ('better quality of care') and control belief 2 ('guidelines have substantially changed their recommendations'). Only control belief 2 remained in the final model (Table 7.13). Model 2 explained 9% of variation in secondary outcome and suggested that GPs who believed more strongly in 'BTS asthma guidelines have substantially changed their recommendation' had smaller prophylactic to bronchodilator cost ratios. The addition of demographic and practice characteristics variables (including those significant in Model 1) did not improve this model.

Efficiency indicator: inhaled corticosteroid cost per ADQ

The indicator was meant to capture efficient use of resources for asthma prescribing. Evidence suggested that it was appropriate to use the cheapest available inhaled corticosteroid when prescribing for asthma (Eccles et al,

2001b). Costs (NIC) were expressed in pounds sterling. ADQs for different preparations were presented in Table 7.1. The distribution of data for this outcome was close to normal (range: 0.11 to 0.75; median=0.32; mean=0.29; SD=0.12).

TABLE 7.14. REGRESSION MODELS FOR EXPLAINING THE VARIATION IN EFFICIENCY INDICATOR OF PRESCRIBING FOR ASTHMA

Dependent variable	β coefficient (CI)	p-value for β	Model's F value (p value)	Model's R square
Model 1 (TPB variables)			6.0 (<0.001)	0.22
Behavioural belief 1: patients' will be healthier	-0.007 (-0.01 to -0.002)	0.005		
Control belief 6: hearing about BTS asthma guidelines	0.003 (0.001 to 0.006)	0.01		
Normative belief 4: PCO prescribing adviser	0.004 (0.001 to 0.008)	0.02		
Control belief 3: under time pressure	0.004 (-0.001 to 0.009)	0.10		
Model 2 (all variables)			5.5 (<0.001)	0.34
Behavioural belief 1: patients' will be healthier	-0.007 (-0.01 to -0.002)	0.007		
Asthma clinic	0.07 (0.02 to 0.12)	0.009		
Practice deprivation score (two categories)	0.05 (0.005 to 0.01)	0.03		
Normative belief 4: PCO prescribing adviser	0.004 (0.00 to 0.008)	0.03		
Control belief 6: hearing about BTS asthma guidelines	0.003 (0.00 to 0.006)	0.04		
Computerised practice	-0.08 (-0.17 to 0.01)	0.07		

Variables excluded from Model 2 because univariate p > 0.2

Consultation length	Age	Years since graduation
Training status	Fund-holding status	Dispensing status
Gender	Senior partner in practice	

Variables excluded from Model 2 after initial inclusion (because p for regression coefficient > 0.2)

Number of WTE GPs in practice	All partners graduated in the UK	Interest in asthma
Practice list size		

At the first stage the TPB direct and indirect variables were assessed for their contribution to explaining variation in efficiency outcome. Indirect and direct subjective norms as well as indirect perceived control met the criteria for inclusion in the model, but the model did not reach statistical significance. Also

self-identity measures were not related to the efficiency outcome. At the next stage belief based TPB variables were considered. Six variables met the minimum criteria, but two (normative beliefs 2 and 3) were excluded from the model because of minimal contribution and presence of multi-collinearity. Model 1 explained 22% of variation in efficiency outcome (Table 7.14). The model suggested GPs believing more strongly that adhering to BTS asthma guidelines improves patients' health were likely to be more efficient prescribers. In contrast those GPs who believed more strongly that they constantly hear about asthma guidelines and those who perceived more strongly that PCO prescribing advisers expected them to follow the BTS guideline were likely to be less efficient prescribers (Table 7.14).

Model 2 was developed by addition of demographic and practice variables to Model 1. All belief based variables that were significantly related to the efficiency outcome in Model 1 kept their statistical significance in the latter model. Model 2 explained 34% of variance in efficiency outcome. It suggested that as well as the conclusions of Model 1, the presence of asthma clinics in practices and high deprivation scores (binary variable based on median score) were associated with less efficient prescribing. As the observed association with binary deprivation scores disappeared when the linear variable (i.e. square root of original score) was used in the model, it should be treated with caution.

7.4. Discussion

This study was probably the first TPB study of GPs' adherence to clinical guidelines and probably the first TPB study that considered both prescribing behavioural intentions as well as GPs' actual prescribing practice. Similar to the findings of previous studies of health professionals (Table 6.2), the TPB proved to be a powerful model for explaining variation in behavioural intention. Up to 47% of variation in reported intentions to follow clinical guidelines' prescribing recommendations were explained in different models (Tables 7.6 and 7.11). TPB elements also demonstrated some ability to explain variation in prescribing outcomes. Belief based TPB variables explained up to 14% of variation in statins prescribing indicators and 22% of variation in asthma prescribing indicators (Tables 7.8, 7.9, 7.13 and 7.14). These latter explanatory powers were achieved, however, through significant deviations from standard methods of analysis TPB items as suggested by Ajzen and others.

7.4.1. Statins prescribing

The study demonstrated that attitude and perceived controls were the main explanatory variables for variation in statins prescribing intentions (Table 7.6). According to this finding, subjective norms were not important elements in GPs' decision making processes when prescribing statins for secondary prevention of CHD. This might have happened because statins prescribing might be more correlated with the GPs beliefs in strength of evidence of effectiveness and organisational and monetary barriers rather than expectations of colleagues and other professionals. It might also be the result of poor performance of the composite scoring approach for subjective norm measurement.

Similar pattern was observed in the analysis of the effective delivery of statins prescribing. No normative belief was related to the outcome, while one behavioural belief (the likelihood of getting quality markers) was linked to the

outcome. A further behavioural belief and one control belief also approached statistical significance (Model 1, Table 7.8). The observed effects of 'getting quality marker' belief were small (small coefficients). There might be different explanations for the finding. It might be the result of statistical artefact or it might not reflect GPs true feelings and beliefs. If the responses were valid representation of GPs beliefs, then the findings suggested that when GPs perceived statins prescribing as likely to result in 'quality markers', they were more likely to follow it. In contrast, other beliefs did not show such effects.

Unlike intention and the effective delivery outcome, subjective norms had significant contribution to efficient prescribing outcome. Different beliefs explained 12% of variation in efficient prescribing. Those GPs who perceived more pressure from PCO advisers for statins prescribing were likely to prescribe less efficiently (Table 7.9). The other interesting finding here was that belief based TPB variables were able to contribute significantly to the model in presence of demographic and organisational factors. Analysis of the efficiency indicator yielded other interesting findings. Senior partners and those working in larger practices were more likely to spend more on statins DDDs. In reverse having CHD clinic or interest in CHD were linked to less costly prescribing (Model 3). More efficient prescribing because of CHD prevention clinics was an addition to their previously shown positive effects on quality of care and death rate in experimental studies (Murchie et al, 2003). Higher reported perceived control was correlated with more efficient prescribing (Model 1). Among the control beliefs the NSF was significantly correlated with the outcome. Interestingly, control belief 6 (statins prescribing may exhaust practice's prescribing budget) was not correlated with the outcome. Among all the variables contributing to Model 3 (Table 7.9), belief items, interest in CHD and clinics were changeable without major revisions in the structure of primary care. Further studies were required to estimate how much cost was involved in changing beliefs and interests.

The analyses suggested practices with higher deprivation scores had higher values of the effective delivery indicator (Model 2, Table 7.8). This was different from previous literature on lower quality of care offered to socially deprived patients (see Chapter 6). One explanation might be that among

patients younger than 60, socially deprived patients who were exempt from prescribing charges might have more incentive to continue their treatment course with statins than those that were marginally better off but found the continuing costs a disincentive. Some limitations should be noted here before making any conclusion. First the variable used in the mode was binary (using the median), because the original variable was positively skewed with many practices having a score of zero. Therefore the finding did not suggest dose response pattern. Second higher prescribing might be the result of higher needs of the patients (i.e. CHD; see Smith et al, 1998b). Third, the data did not demonstrate whether it was the deprived patients who were receiving these statins or other patients registered with those practices (i.e. ecological fallacy; see Morgenstern, 1995). Answering these queries with PACT data was not possible and required further research including patient data. The analysis however did not indicate that people living in socially deprived areas were receiving lower quality care.

7.4.2. Prescribing for asthma

Indirect subjective norms and direct perceived control variables were able to explain 47% of variation in intentions to prescribe in accordance with the BTS asthma guidelines (7.11). The absence of any significant contribution from attitudes was unexpected as attitudes were generally the most powerful predictors of intentions in TPB studies (Lambert et al, 1997; Levin, 1999; Walker et al, 2001; Watson and Myers, 2001; Puffer and Rashidian, 2004). TPB variables had no contribution to the primary effective delivery outcome of prescribing for asthma (corticosteroid DDD per weighted STAR-PU). It was not possible to ascertain whether this showed genuine lack of relationship between psychological constructs of the TPB and the outcome, or difficulties in measuring a valid indicator of asthma prescribing from PACT data (7.2.5).

A significant relationship was observed between one control belief and secondary effective delivery outcome (corticosteroid cost to β-agonist cost ratio) (Table 7.13). The analysis suggested those GPs that felt more strongly that

'BTS asthma guidelines recommendations have substantially changed' were likely to prescribe less in accordance with the guidelines' recommendations. This was the only variable with significant contribution to the model. This finding was consistent with the qualitative study (Chapter 4) which suggested the notion of change was an important factor for guideline implementation, and with the analysis of asthma prescribing intentions (7.3.8). This control belief did not feature in the models explaining efficiency indicator (Table 7.14). The belief and its potential effect on prescribing might be a response to some of the changes in the BTS asthma guidelines recommendations through the years (British Thoracic Society, 1993a; British Thoracic Society and National Asthma Campaign, 1997a; British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2003). However, again it was difficult to establish whether this association was due to statistical artefact or genuine relationship between the belief and the outcome. It would be interesting to know what level of change in recommendations could cause enough discontent in GPs to hamper guideline implementation. Such studies might be difficult with real life scenarios, as there were a lot of confounding factors. It might be possible, however, to conduct a study using guideline simulations and record the reaction of GPs to the change in recommendations. This might help the development and implementation of clinical guidelines in future. On the other hand, asthma care in British primary care was often offered by practice nurses under GP supervision. Therefore, changes in recommendations might not affect GPs in terms of time spent on patient education or losing face in front of patients (e.g. because GPs changed their recommendations). Similar explanation might also apply to CHD preventive care. The qualitative study (Chapter 4) suggested that statins were usually prescribed by GPs while they were consulting patients, while asthma prescribing happened as a result of assessments conducted by practice nurses.

Model 1 in Table 7.13 suggested having 'asthma clinic' in practice was correlated with lower quality of asthma prescribing. This finding was puzzling, since case studies demonstrated that asthma clinics were linked to many aspects of better asthma care (Dickinson et al, 1998). Also clinical guidelines emphasised the importance of providing asthma clinics in general practice (Scottish Intercollegiate Guidelines Network, 1998). On the other hand a randomised trial of using specialist nurses to support practice nurses in asthma

care failed to improve patient outcomes (Premaratne et al, 1999). An explanation for the survey's finding could be that asthma clinics had improved the management of patients in a way that short-acting β_2-agonist bronchodilator treatment was adequate for bigger proportion of patients. As a result, less amount of inhaled corticosteroid were prescribed per weighted STAR-PU.

Similar to the analysis of efficiency indicator of statins prescribing, TPB belief based variables were powerful predictors of variation in the efficiency outcome of asthma prescribing (Table 7.14). Three behavioural beliefs explained up to 22% of variation in the outcome and retained their significant contribution to the model in presence of demographic and practice variables (Table 7.14). Again, as for the statins efficiency outcome, perceiving more pressure from PCO prescribing advisers was linked to less efficient prescribing for asthma. Also the analysis suggested that GPs who heard more about the BTS asthma guideline were less likely to prescribe efficiently. Among all demographic and practice characteristics, three had significant contributions to the model. Computerised practices were likely to prescribe more efficiently, while those with asthma clinics or higher deprivation scores seemed to be less efficient in their prescribing of inhaled corticosteroids (Table 7.14). It was known that asthma clinics might result in increase in asthma treatment costs (Dickinson et al, 1998), but there should be no reason for it to reduce the efficiency in choice of inhaled corticosteroids. A previous study suggested there could be interaction effect between practice deprivation and the corticosteroid to bronchodilator ratio (Shelley et al, 1996), in which in more deprived areas, the ratio was less 'preventive' of asthmatic admission to hospital. In other words higher ratios did not reflect themselves in reduction in admission rates. This might also be the result of higher levels of need in deprived areas (Salamzadeh et al, 2003). The caveats discussed in the previous section on statins, apply here as well.

7.4.3. Self identity, expressed interest and practice characteristics

Self identity measures did not contribute to models intended to explain variation in prescribing intention. Interestingly, GPs who perceived themselves as evidence-based practitioners were likely to prescribe more statins per weighted prescribing unit (Table 7.8). This observed effect turned non-significant in presence of demographic and practice characteristics. No other contribution to multi-variate models was observed from the evidence-based practitioner self-perception or the patient-centred practitioner self-perception.

Expressed interest in CHD was correlated with more efficient statins prescribing (Table 7.8) but not more effective prescribing (7.9). The analyses did not identify any links between expressed interest in asthma and prescribing outcomes. This was different from other studies (Jones, 1992), which concluded that interest in asthma reduced prescribing costs. Campbell et al (2001a) concluded that the training status of the practice was not a significant predictor of quality of care. Training status did not feature in any of the multivariate models of prescribing outcome indicators. No practice or service characteristic was linked in multivariate analysis with the effective delivery indicator of statins prescribing. Only having CHD clinic was correlated with efficient prescribing of statins (Table 7.9). Computerised practices presented themselves as effective prescribers for asthma (Table 7.13). Rather surprisingly, GPs who reported having asthma clinics prescribed less efficiently and less effectively as measured by the indicators (7.4.2; also Tables 7.13 and 7.14).

Previous studies identified strong relationships between practice booking intervals and the quality of chronic disease management. One large survey concluded that the scores for quality of care for asthma were 67% higher in practices with 10-minute booking intervals for routine consultations than in practices with 5-minute intervals (Campbell et al, 2001a). The quality scores were calculated using a set of criteria including prescribing. No such relationship was identified in the surveys' analyses. This might be due to the lack of statistical power to capture these potential differences in the surveys.

Only one GP in the statins survey and one in the asthma survey reported 5-minute booking intervals.

7.4.4. Valid measurement of prescribing

TPB assumes that behaviours were under volitional control, and therefore attitude towards *behaviour*, and not attitude towards *target*, predicted behavioural intention and behaviour (Montano and Kasprzyk, 2002). The difference between these two types of attitudes was discussed in Chapter 3 (3.2.6). It is common to assume physicians' behaviour as volitional and the result of careful decision making. Clinical preferences and choices of physicians were referred to as 'individual clinical policies' implying deliberate decision making (Lomas and Haynes, 1988). However, variations in physicians' behaviour might not be volitional in all circumstances. The review of the impact of computerised reminder systems demonstrates well that some variations were due to errors of omission (Soumerai et al, 1989). A long term study demonstrated that as soon as computerised reminders were stopped the rate of antibiotic prescribing for streptococcal pharyngitis returned to lower pre-intervention levels (Barnett et al, 1978). In that study all participating physicians had agreed to take part in the reminder system and considered antibiotic prescribing for streptococcal pharyngitis as clinically important topic. Variation in prescribing might be down to simple (or complicated) organisational variables despite individual physicians' clinical policy. In theory, TPB was capable of capturing these variations under the perceived control concept. Suppose a physician was asked about his or her perceived ability to prescribe appropriate antibiotics for all cases of streptococcal pharyngitis requiring the treatment. It could be assumed that when the reminder system was not in place, the physician would express less ability to perform the behaviour. In reality, this assumption depended on the physician being aware they had omitted some cases that required the treatment. Unless the physician was aware of the omission, the TPB's perceived control concept might not capture the variation.

Prescribing outcomes might not be accurate reflections of GP intention and activity, rather the effects of other clinicians' behaviour. Practice nurses were actively involved in the management of chronic diseases and their preferences could affect GP prescribing. Nurses usually had the authority to decrease or increase the dosage of prescribed treatments, although they consulted the GPs for the initiation of new medicines (Dickinson et al, 1998). A proactive practice manager who pursued the identification of patients requiring secondary prevention or those with asthmatic symptoms might contribute to higher levels of prescribing, without direct effects on the beliefs of the GPs. GPs in partnership practices might sign the repeat prescription orders for patients under the care of their GP colleagues, without being directly involved in decision making for the care of the patient. Also some medicines assessed in this study could be consultant initiated. For example, if the local consultant prescribed expensive products such as fluticasone propionate more frequently, the GP prescribing might appear less efficient as a result. Or if a local hospital was proactive in ensuring all coronary artery patients were discharged on statins this would improve the performance of local GPs without necessarily affecting their views or beliefs.

Habits and routines contribute to clinician behaviour. GP prescribing for certain conditions might be due to *routines* more than to decisions. An analytical study of GPs' cognitive processes while prescribing (antibiotics) for UTI or stomach complaints concluded that up to forty per cent of prescribing decisions were due to habits (Denig et al, 2002). They also concluded that habits resulted in optimal or suboptimal prescribing for patients. If that was the case, then attitude towards behaviour were less influential in affecting the behaviour. Despite Denig et al (2002), the term 'routine' was deliberately used in here instead of 'habit' to acknowledge that prescribing required information processing. Behaviours such as prescribing (for given conditions) that were not completely volitional, but were not as automatic as habits, better to be referred to as routines (Norman and Conner, 1996). The spontaneous processing model accounted for behaviours which were not mediated by decisions following careful considerations. This model suggested that in these circumstances, attitude towards *object* or target might explain behaviour better that attitude towards *behaviour*. The model argued that attitude towards object could be

activated in response to certain cues and lead to change in behaviour (Norman and Conner, 1996). The model also suggested that deliberate consideration of behaviour happened more in novel conditions, or for new problems, than in familiar conditions (Norman and Conner, 1996). The implications of this argument for guideline implementation was that careful and volitional assessment of patients followed by deliberate implementation of guideline recommendations was more likely to happen for new patients or diseases, or new presentations of chronically ill patients. GPs might prescribe many drugs because of routines (similar patients, protocols) or previous decisions (repeat prescription in chronic diseases). It might also imply that changing prescribing for treatment of chronic diseases or common diseases (with chance of seeing many seemingly similar patients) were more difficult than for other diseases.

It was important to distinguish behavioural attempts from actual behavioural performance. For example GP might intend to prescribe statins for secondary prevention of CHD in accordance with the guideline. The doctor offers the prescription to the patient. The patient may refuse to take the medicine in two ways. He or she may refuse to accept and as a result the GP decides to offer an alternative medicine, non-medicinal treatment or none for the time being. Second, the patient may receive the prescription, but decides not to go to the pharmacy and take the medicine. In both of these scenarios, there has been a behavioural attempt, but no behavioural activity. The outcome measure in the study is only capable of capturing actual behaviour outcomes through drug dispensing information (PACT data). Note that the second scenario is not, even in theory, captured by the perceived control element of TPB. The positive behaviour of the GP has not resulted in prescribing because of the patient's decision. The GP has no control over the patient's decision and may not become aware of it. This is another limitation of studying clinicians' behaviour. In the context of the TPB, both behavioural attempts and actual performances form what called behavioural achievement (Ajzen, 1991). However, behavioural attempts are very difficult to measure in prescribing context. Also, for example, if the patient who is supposedly on the treatment does not take the drug regularly, it will affect its need for the medicine and therefore the amount dispensed through repeat prescription. All these scenarios suggested that future TPB studies would do better if considered doctors'

prescriptions as the outcome rather than dispensed medicine. Future studies should also focus on new patients or new prescriptions and attempt to separate repeat prescriptions.

It has been argued that patient variables such as culture, economic and social circumstances and patients' perception of the outcome were more important variables of variation in the provision of health care than doctors' recommendation (Greer et al, 2002). Crude social variables did not capture this variation. The study did not measure patient characteristics. Indeed it was difficult to measure these in valid and reliable ways. Nevertheless this might have contributed as one of the factors that resulted in incomplete capture of variance in prescribing indicators.

7.4.5. Specific recommendations versus guideline as a whole

The surveys assessed only some of the messages provided by the clinical guidelines. Clinical guidelines very often give non-prescribing advice as well as prescribing advice. The surveys focused on specific prescribing issues. Therefore the totality of the guidelines' advice was not considered. Focused approach is common in guideline implementation studies. Two investigators systematically reviewed several clinical guidelines and non-interventional and interventional studies of guideline implementation (Chrischilles and Gondek, 1997). Among twenty-four non-interventional studies they found only one study that considered all the prescribing advice offered in the studied guideline. They concluded that the majority of guidelines did not provide detailed advice about drug regimen issues that were required for the evaluation of prescribing. Similarly, another systematic review identified only one study that '*looked at a broader range of quality of care issues, including inappropriate choice of therapy, under-dosage and over-dosage, scheduling, drug-drug interactions, therapeutic duplication, allergies and the omission of necessary drug therapies*' (Anderson and Lexchin, 1996). This might be because of difficulties in measuring appropriate outcomes. On the other hand it might be due to the limited information provided in many clinical guidelines. Very often there is not

enough evidence to support detailed recommendations. Even if the guidelines provided detailed advice, it would be difficult to link recommendations to measurable outcomes. A review of nineteen primary care clinical guidelines concluded that it was not possible to link most of the guidelines' recommendations to measurable outcomes (Balas et al, 1994).

TPB has stronger theoretical footing when applied to behaviour rather than consequent outcome and to single behaviour rather than multiple behaviours (Bagozzi et al, 1989; Conner and Sparks, 1996). These issues impose limitations on studies using TPB to examine adherence to guidelines. Adherence to guideline involves multiple behaviours. It might include correct diagnosis and diagnostic investigations, proper consultation with the patients, drug and non-drug prescriptions, management of adverse events and follow-up. Thus the goal of adherence is achieved only if most or all of those behaviours are performed. To assess the feasibility of the TPB in explaining variation in guideline implementation required focus on limited areas of guideline recommendations. It is possible, at least in theory, to capture all recommendations made by guidelines, but it would require questionnaires of several pages long that would reduce the chances of getting meaningful responses from participating GPs.

7.4.6. Study design and data sources

Prospective designs are appropriate for TPB studies. Behavioural intention and its determinants are measured at one point of time and behaviour at a later stage (Conner and Sparks, 1996; Montano and Kasprzyk, 2002). Simultaneous measurement of behaviour with other elements of the model (cross-sectional design) is not recommended and might provide poor and unreliable measures of past behaviour (Montano and Kasprzyk, 2002). The analysis of the surveys' prescribing outcomes was essentially prospective. The responses to the questionnaires were collected in early 2002, while the prescribing data was for April to June 2002. Another important feature of the design was the attempt to capture documented behaviour from PACT data instead of relying on reported

behaviour, which was common in TPB studies. Self reported behaviour has higher correlation with attitudes and intentions and might be subject biases (Armitage and Conner, 1999b; Armitage and Conner, 2001).

The primary outcome measures used in the surveys were calculated from PACT data. PACT was not linked to patient records; hence it was not possible to assess the appropriateness of prescribing for individual patients. Using PACT data, the investigator was incapable of establishing whether those patients who required drug therapy based on their diagnosis were given the drug appropriately, and whether those who should not have received a drug were given it. Prescribing units captured only part of these variations. GP lists were inflated by about six per cent (National PCT Database, 2004), because movements (of young adults), emigrations and deaths are not quickly updated. Therefore the practice demographic data and hence the STAR-PUs might not reflect the truth. It was not possible to establish the level of inflation for different practices. This was likely to introduce error into the analyses of effective delivery outcomes and reduce the significance of regression coefficients.

The arguments in favour of or against the use of different prescribing indicators were presented in the methods section (7.2.5). In addition, a consensus development study assessed the validity of prescribing outcomes (Campbell et al, 2000c). The study used Delphi method (Adler and Ziglio, 1996) to identify valid prescribing indicators of quality and efficiency of prescribing based on PACT data. Out of 41 indicators that were assessed for both quality and efficiency, only 12 (out of 41×2=82) were considered valid by participants. The 'valid' indicators did not include indicators of prescribing for asthma or of statins. However, an indicator for asthma measuring cost per DDD of inhaled corticosteroids scored reasonably well for efficiency (7 out of 9). Another asthma indicator, ratio of inhaled corticosteroids DDD to inhaled corticosteroids STAR-PU scored 6 out of 9 as a quality indicator. An indicator 'items of lipid lowering drugs per patients aged 45-75' also scored 6. The only statins prescribing indicator considered in the study was 'no of items for statins per 1000 patients'. This indicator understandably scored low for both prescribing quality and efficiency.

The study design could have been improved by recruiting GPs from practice linked to databases such as General Practice Research Database (GPRD), Doctors Independent Network (DIN) or MediPlus database (Lawrenson et al, 1999; Hansell et al, 1999). The advantages of using these resources were that GPs prescribing data were linked to individual patients and their recorded diagnosis (see Appendix VIII). Hence it would have been possible to reduce variation in prescribing outcomes occurred because of different sources of inaccuracy. The main disadvantage of the GPRD database was that it might not be representative of the GP population. However, if high response rate was achievable from the GPs linked to this database, the disadvantage might be ignored. It was not possible to use GPRD or similar resources because it required financial resources that were not available to this PhD study.

7.4.7. Response rate, questionnaire design, optimal scaling and TPB analysis

In recent years the response rate of health professionals (in particular GPs) to research surveys has been falling (Kaner et al, 1998; Barclay et al, 2002). Low response rates may threaten the validity of the findings of surveys. The detailed non-response analysis identified differences between the respondents and non-respondents, but most did not reach statistical significance (Table 7.3). There was significant difference between the two surveys, as more GPs responded to the statins surveys. Apart from this the only other clear difference between the two groups was the higher response rate from training practices. This was statistically significant in overall analysis and for the statins survey but not the asthma survey. If the lack of significant relationship between the training status of the practice and responding to the asthma survey was not statistical artefact it might suggest that GP trainers no longer found asthma an interesting topic. The response rate from those GPs who were working in practices in which all GPs were qualified in the UK (regardless of their ethnic origins) was higher than others. 59% of the sampled GPs were working in such practices, but 65% of the responding GPs were from them (p=0.06; Table 7.3).

Despite strenuous efforts the response rates were poor. The efforts included two reminders, reducing the number of questions in the main surveys, better quality print for the questionnaires, pre-paid reply envelopes, individually addressed and hand-signed covering letters every time, changing covering letter content for the first and second reminders and the provision of a lottery for respondents. A factor that might have contributed to low response rate was the request for access to the GPs' PACT data. Anecdotal feedback from respondents suggested that some found asking for access to PACT data a reason for not returning the questionnaire. The other potentially important factor was the length of the questionnaires. Shorter questionnaire might have led to a better response rate. Shortening the questionnaires should be done without excluding important elements of the TPB. In the light of that, it might be possible to reduce the number of items in direct attitude and intention scales. This, however, would not shorten the questionnaire dramatically. Most of the questionnaire items were belief based and the findings strongly supported their inclusion in TPB studies. Therefore, questionnaire length would remain the limiting factor for TPB studies of health professionals. Another solution for reducing the number of items in the questionnaire was to drop value statements from belief based items. This could shorten the questionnaire substantially. In a recently published study of primary care physicians, the investigators dropped value items; instead they multiplied positive beliefs by an average positive value and multiplied negative beliefs by an average negative value (Montano and Kasprzyk, 2002). They did this to shorten the questionnaire. Shortening the questionnaires in this way has theoretical drawbacks, as one expects different beliefs to elicit different values from different physicians. The validity of solutions like this also depends on finding solutions to the more serious limitation of TPB studies, namely the inability of the theory to suggest optimal scaling approach for the questionnaire items (7.3.2).

It was recently suggested that the analysis of composite scales as suggested by Ajzen and others introduced biases in the findings and made interpretation difficult (French and Hankins, 2003). They considered this to be a problem of modelling the interaction between expectancy and value items, rather than the theory. One proposed solution was to put expectancy and value

items in the regression model and then add composite measure (i.e. interaction) to see how much more of the attitude's variance was explained by the interaction. But the authors argued this was not psychologically sound (Hankins et al, 2000; French and Hankins, 2003). Future studies might use statistical optimisation to identify the best scaling approach and composition strategy for the analysis of the TPB items (Box and Cox, 1964). Such optimisation should be based on datasets of a few different studies of health professionals, ideally all with prospective measures of outcome behaviour

Previous studies proposed that subjective norm was the weakest explanatory variable of intention (Godin and Kok, 1996; Armitage and Conner, 2001). Hence the power calculation for the surveys was based on the relationship between subjective norm and intention (Chapter 6). The results of the surveys did not support that assumption. Subjective norms performed as well as other elements of the TPB in the models. More importantly, one normative belief remained significant contributor to variation of the outcome for both statins and asthma surveys. This might be the result of careful attention to the measurement of subjective norm and normative beliefs in the surveys (Armitage and Conner, 2001). Careful attention was given to the optimal scaling of TPB indirect measures. Many previous studies shunned this important consideration (e.g. Walker et al, 2001).

It was argued that to test a theory, it was not always suitable to conduct empirical research (Smedslund, 2000). Smedslund proposed that the distinction between a priori positions and empirical questions were not always acknowledged in psychological research. He argued that for a priori positions 'rational evaluation' was the way of testing the position and not empirical research. This could have important implications for the way the TPB was evaluated. Because of similarities in item wording it seemed plausible to expect high correlation between the items, e.g. between intention and reported past behaviour, or between attitude and intention. Some of the relationships observed in the analysis of the TPB were logical. Davidson (1998) documented similar criticisms over the transtheoretical model – that high internal reliabilities observed in empirical studies were the mere results of similarities between the meanings of the items (semantic overlap). Where there was semantic overlap

between scale items, the psychometric characteristics of the scale could be misleading (Davidson, 1998). Also good internal reliability estimates might have happened because of the way items were presented. Likewise, if items had similar meanings, they would share specific variance and common factor variance. Similarity of item wordings for proximal constructs of the model could have resulted in good psychometric capabilities only because of inter-item semantic overlaps. Nonetheless there were elements in the TPB that warranted empirical testing. Among those were the relative strengths of different elements of the theory in explaining variation in intention. The criticism applied most to explaining variation in intention using direct TPB variables. The criticism might not apply to the models that explained intention by indirect measures or the models that explained actual behaviour (e.g. prescribing). The TPB belief based constructs performed reasonably well in the surveys' models where intention was the predicted variable, as well as in the models with prescribing as predicted variable; however in the latter group the belief based items were used as individual items.

Measurement of reported past behaviour in the surveys followed Ajzen's and others' recommendations. This was not ideal, especially as the format of reported behaviour question became similar to the format of the intention question (Armitage and Conner, 2001). An alternative approach would have been to use written case simulations or vignettes and ask the GPs to report their behaviour in response to the vignette. Ideally, minimum of two case simulations should have been used to capture different aspects of past behaviour. However, addition of even one vignette could have resulted in significantly longer questionnaire. It should be noted that in the analysis of the surveys, the reported past behaviour was used only in complementary analyses and it did not contribute to the variation in the prescribing indicators.

Inclusion of social desirability questionnaire might have increased the validity of the findings (Crowne and Marlowe, 1960). Following clinical guidelines' recommendations for secondary prevention of CHD and treatment of asthma were 'socially desirable' behaviours. Hence the respondents might have unintentionally ('social desirability') or intentionally ('faking good') expressed themselves more positive towards the behaviours than they really were

(Streiner and Norman, 2003). However, inclusion of further questions in the questionnaire was not feasible. The questionnaires were already long enough to discourage some GPs from responding. Longer questionnaires might have resulted in lower response rates. Also previous studies suggested that social desirability had minimal impact on TPB models (Armitage and Conner, 1999a).

Another approach would have been to ask about different set of behavioural beliefs in the questionnaires. The actual behavioural beliefs were targeted towards the ultimate outcomes of evidence-based prescribing for patient outcome and improved quality of life. Alternatively, these items could have targeted immediate behavioural outcomes, e.g. reducing number of asthmatic attacks, re-admissions, blood cholesterol level, risk of cardiac events or even number of visits to the practice. Using these items might have resulted in bigger correlation coefficients between beliefs and intention or prescribing indicators. On the other hand this approach might have reduced the correlations as these specific outcomes might not necessarily correlate with the goal of achieving better quality of care for patients. Also specific questions might not necessarily correlate with the totality of following clinical guideline recommendations, as they might relate to individual patients or to specific ways of prescribing. These might alter from time to time even for individual patients. Nonetheless specific outcomes did not demonstrate themselves in the qualitative interviews as salient beliefs for GPs when they thought of implementing guidelines.

Future studies might also assess the TPB in the context of diseases with more controversies surrounding their management. These studies could contrast different beliefs against each other. For example, GPs might have strong behavioural beliefs in support of a guideline, but have negative normative beliefs since their colleagues disapprove of the recommendation.

7.4.8. Can the TPB improve understanding of GP prescribing?

The TPB surveys targeted prescribing for two specific conditions. Therefore, the results were not included into the taxonomy of interventions presented in Table

4.2. In order to use TPB surveys to explain variations observed in the effectiveness of interventions, TPB studies should be conducted along-side experimental studies. The surveys' results demonstrated that control beliefs (whether GPs perceived themselves to be able to undertake the behaviour) and normative beliefs (whether GPs perceived important others approved or expected the behaviour) were at least as important as behavioural beliefs (whether the behaviour resulted in suitable outcomes) in behaviour formation. This has important implications for clinical guidelines development and implementation. Although the methods of developing valid guidelines have improved in recent years, less is known on how to incorporate the views of target practitioners into the guidelines. Recent studies demonstrated the role that physicians' beliefs played in their behaviour, including their prescribing. A nested observational study of 800 patients in five British general practices concluded that *perceived* medical needs of patients was the main predictor of GP prescribing (Little et al, 2004). They also observed that doctors were more likely to prescribe in response to patients' pressure if they perceived the pressure. By interviewing patients as well, they found that GPs' perceived pressure was independent of real expectations of the patients.

The theory performed well in explaining variation in GPs stated intentions to adhere to guidelines prescribing outcomes. This was similar to previous studies of health professionals (Table 6.2) and health behaviour in general (Ajzen, 1991; Conner and Sparks, 1996; Armitage and Conner, 2001). Although the observed ability of the direct variables (attitude, subjective norm and perceived control) in explaining variation in intention could be explained in part by semantic overlap (7.4.7), the indirect variables had high correlations with the intention, and in that sense the TPB seemed internally valid.

Attempts were made to identify optimal scaling approaches for belief based items for the calculation of expectancy-value composite scores. When this was possible (behavioural beliefs), different scaling approaches did not make a lot of difference to the outcome of the analysis and it was possible to identify a relevant scaling approach. For control beliefs, the study was not able to suggest any optimal scaling approach. It should be noted that although it was possible to support empirically previously suggested scaling approach for

behavioural beliefs (Ajzen, 2002b) and there was agreement on how to scale normative beliefs (Conner and Sparks, 1996; Ajzen, 2002b), these did not tackle the main problem of 'sums of products' calculations and scaling, which was the difficulty of establishing the location of the natural zero on the scales (French and Hankins, 2003). Further methodological studies were required to investigate the options available (7.4.7). Future works should consider the potential for variable scaling of items within beliefs as well as between beliefs and also expectancy statements.

TPB elements explained parts of variation in prescribing indicators. The consistency of the findings for one normative belief (PCO prescribing advisers) for both statins and asthma survey supported the basic assertion of the TPB that beliefs influenced the outcome. However, there was no significant relationship between intention and prescribing indicators. The TPB suggested that the behaviour was distal to the formation of beliefs; and beliefs influenced behaviour through the intention. The TPB also suggested perceived controls moderated the effects of intention on behaviour. Control beliefs as well as behavioural beliefs inconsistently featured in models explaining variation in prescribing indicators. Therefore, it was not possible to claim the findings supported the direct effects of control beliefs on behaviour, because other groups of beliefs showed similar capabilities. Nonetheless, the findings were against the theory of reasoned action assertion that behaviours were conducted through the formation of behavioural intentions. The existence of direct links between different beliefs and prescribing indicators opened the issue of another difficulty in testing the TPB. In reality there are certain degrees of overlaps and correlations between different beliefs (Ajzen, 1991). For example normative beliefs (e.g. GP's perceived belief that colleagues expected him or her to follow guideline's recommendations) could be linked to behavioural and control beliefs. The GP's colleagues could affect his or her behavioural beliefs (e.g. whether guideline recommendations resulted in improved patient health) and his or her control beliefs (e.g. whether following guideline recommendation resulted in waste of practice resources). Therefore although it was sensible to look for identification of beliefs in each category, they were not mutually exclusive.

The contribution of beliefs in explaining variations in prescribing outcomes was encouraging and improved the understanding of GP prescribing. The stability of these findings was not assessed as required larger sample sizes. Also there were differences in the explanatory contributions of different beliefs to the asthma and statins models. This could have been caused because of instability in the models. However, it should be noted that these were correlational models. When two regression models have different explanatory variables and there are correlations between the explanatory variables included in the models, differences in explanatory variables might not be because of instability but because of those correlations (Russell and Gregson, 1981). There were also differences in the models prescribing outcomes. Different outcome variables (for asthma and statins) could be explained by different explanatory variables. Another difference between the outcome measures was that the efficient prescribing indicators were more valid than the effective delivery indicators. The differences might have also been caused by non-response bias, so that the respondents to the asthma survey were different from the respondents to the statins survey.

The analysis of prescribing indicators suggested that beliefs (e.g. normative beliefs) were better predictors of behaviour than the composite scores for their corresponding higher level construct (e.g. indirect subjective norm). The finding that some beliefs but not others significantly contributed to the models was not against one of the underlying principles of the TPB, namely that intentions and behaviours were based on salient beliefs. But it again questioned the usefulness and reliability of composite scoring approaches recommended by the TPB.

The findings suggested that the TPB helped understanding of GPs prescribing behaviour and their intentions to implement a clinical guideline. TPB also featured well in the design and analysis of the qualitative study (Chapter 4). TPB models should be tested in future alongside RCTs where prospective collection of valid outcomes could minimise the caveats and help to assess the theory in controlled situation. It will also help to test the assumption of causality in the TPB (i.e. change in beliefs ultimately results in behaviour change).

We live in an age of mass loquacity. We are all writing it or at any rate talking it: the memoir, the apologia, the cv, the cri de coeur. Nothing, for now, can compete with experience—so unanswerably authentic, and so liberally and democratically dispensed. Experience is the only thing we share equally, and everyone senses this.

Martin Amis, *All from experience*, (Amis, 2000). Quoted from *the BMJ* 323, 162

Knowing is not enough, we must apply; willing is not enough, we must do

Goethe

Chapter 8. Conclusions and implications

8.1. As a prologue: revisiting the complexities

Provider (including physician) behaviour change can be analogised to a maze. Change or improvement, as targets, are achieved when one gets through the maze and reaches to the anticipated exit. However, each policy-maker or practitioner knows only part of the way through this maze. It is difficult to avoid all the dead-ends or predict the time required to go through the maze. A graphical image from the world of arts might resemble the difficulties of behaviour change. In the film "a chump at Oxford (1940)" played by two acclaimed comedians Stan Laurel and Oliver Hardy, they were lost in a maze at the University of Oxford. A student played the role of a ghost that offered them handkerchiefs or cigars by his 'ghost hands' and also played tricks. The 'ghost hands' are invisible barriers or facilitators for behaviour change initiatives. It is extremely difficult, perhaps impossible, to devise in detail a success plan for behaviour change that can be reliably applied to different settings and different occasions. There are players ('other factors') in provider behaviour maze that

are either unknown or have been ignored. These players can have detrimental effects on the outcome of the intervention.

The other problem is that 'provider behaviour maze' has several exits, not all favourable. That is to say the success is not always known. For many interventions, the available evidence at any given time does not provide unchallenged authority for the policy-maker or researcher to say confidently what outcomes can be expected if the intervention is to be implemented. This is truer for external motivation interventions (e.g. financial incentives) or non-voluntary interventions (e.g. rules and obligations) (Table 4.2). Given that no single theoretical perspective (Chapter 3) or empirical initiative (Chapters 2 and 4) is powerful enough to devise all the processes, barriers of progress or routes to success, then it is essential to acknowledge the complexity of provider behaviour and 'proceed with care' (Black, 2001).

Although this thesis was focused on evidence-based practice, it acknowledged that '*there is more to primary care than the use of evidence-based interventions*' (McColl et al, 1998, p 1355). GP capabilities are not limited to the understanding of and the willingness to practise based on the best available evidence. GPs require equipping themselves with consultation skills, team playing skills, practice business management skills, effective communication (with local authorities, colleagues and community) and ways of promoting access to primary care (McColl et al, 1998). Some argue that GPs '*take care of individuals in a society, irrespective of the patient's type of disease or other personal and social characteristics ...*' (Olesen, 2000, p355).

Very often clinical guidelines aim to reduce variation in clinical practice. Not all variation in practice is bad. Variation can be due to valid reasons, including individual patients' needs (Ashford et al, 1999) as well as ambiguity of the best available evidence in demonstrating any advantage for one clinical practice over the other (Burgers and van Everdingen, 2004; Rashidian, 2004d). Hence it is acknowledged that '*not all change is necessarily good and if general practitioners resist changing their behaviour this may be for valid reasons*' (Horder et al, 1986, p 520).

8.2. Principal findings

The thesis was developed based on the findings of three linked studies using different methodologies and focus points in order to enhance the understanding of clinical guideline implementation in primary care. The first study was the overview of systematic reviews (Chapter 2) supported by the selective review of theories and models of behaviour change (Chapter 3). The second study was the qualitative study (Chapters 4 and 5). It was designed and analysed with the help of theoretical insights from theory of planned behaviour (TPB) and the findings of the reviews. Then the qualitative study contributed to the planning of the third study, which consisted of two TPB surveys of GPs (Chapters 6 and 7). The surveys directly assessed the merits of TPB for understanding GP prescribing.

The overview of systematic reviews aimed to identify the effective methods of improving primary care prescribing. Thirty-three systematic reviews were included. The study developed an updated taxonomy of behaviour change interventions for primary care prescribing. It also demonstrated the evidence or lack of evidence for the effectiveness of twenty-three interventions identified from the literature (Table 2.1). The overview observed that all the studies using interactive educational activities (2.3.3) had resulted in positive changes in prescribing. The only other intervention that came close to this was educational outreach visits (2.3.8). However, the effectiveness of this intervention was challenged by more recent evidence not included in the systematic reviews (2.4.1). At the other end of the spectrum, there was 'evidence of no effect' for didactic educational activities (2.3.3) and the substitution model of inter-professional shared care (2.3.7). Evidence for other interventions was mixed, while many had not been assessed in rigorous studies (no evidence of effect).

The overview resulted in important observations. There were several occasions in which the systematic reviewers inaccurately or selectively reported, or unsafely interpreted the results of primary studies. The variation in quality of systematic reviews was known from previous overviews, but this was the first time the inaccuracies were discussed in detail (2.4.2). The overview

also demonstrated that the infamous dichotomy depicting multi-faceted interventions of behaviour change as effective and simple single interventions such as 'mailed printed educational materials' as ineffective was short-sighted (see 2.3.4 and 2.3.16). The findings advocated that the reports of multi-faceted interventions should include the details of the interventions and their intensity. Otherwise, generic messages indicating that multi-faceted interventions were more effective might result in inefficient use of meagre health care resources; especially as some multi-faceted interventions were not more effective than single interventions (2.3.16). On the other hand some interventions with small effects were cost-effective in improving quality of care (2.3.4; see also Mason et al, 2001). The overview also found a few robust studies originated from low and middle-income countries. This was a welcome improvement over what Bero et al observed in 1998. The wider contribution of investigators from around the globe was complementary and increased the understanding of the interventions as well as the diversity of the messages. Still, there were many countries with no studies included in the systematic reviews.

The selective review of theories and models of behaviour change summarised several important theories of behaviour change. It helped identifying theoretical insights that different academic disciplines offered to provider behaviour change (Chapter 3). The review of theories put TPB in the context, especially as TPB shared some of its concepts with other theories. The review also summarised the theoretical justifications for different interventions. A selection of theories that were likely to explain variations observed in the effectiveness of different interventions was listed in the taxonomy of interventions (Table 3.3). The observed effectiveness or lack of effectiveness of some interventions was partially explained by the theories. For example theoretical reasoning had envisaged the ineffectiveness of didactic educational activities and the effectiveness of interactive education. This success was probably due to the fact that educational interventions were in use for several decades and the theories had progressed alongside the development of the interventions. More often, the theories justified certain interventions, while there was not enough evidence to support or reject the theories' claims. Clear examples of such interventions were 'local opinion leaders', 'peer review', 'inter-professional education', 'continuous quality improvement', 'practice support' and

'rules and obligations' (Table 3.3). The systematic reviews included little evidence for the effectiveness of interventions based on certain theories, e.g. management theories. Also interventions with indirect theoretical backing resulted in mixed outcomes including unwanted consequences. For example financial incentives resulted in *significant reduction* in repeat prescription rates (Krasnik et al, 1990). Overt deviations like this were exceptional. In that sense, provider behaviour change did not seem very different from clinical care, where not all treatments devised based on physio-pathological arguments were later supported by evidence.

The qualitative study identified seven important themes and thirty sub-themes for implementing clinical guidelines for primary care prescribing (Tables 4.1 and 4.4). The themes were 'credibility of content', 'credibility of source', 'presentation', 'influential people', 'organisational factors', 'disease characteristics' and 'dissemination strategy'. The study highlighted some important concepts. One was the importance of understanding what GPs meant by the term 'evidence'. 'Evidence' was not necessarily referring to the quality and validity of the study that the recommendation was based on. Rather GPs considered a set of different criteria when referred to evidence. Similarly the participants challenged another common element in evidence-based guidelines: 'change' in recommendations. The qualitative study also contributed to the further development of the taxonomy (Table 4.2). The findings of the qualitative study were compared with the findings of the overview of systematic reviews (Table 4.3). The findings were congruent in several occasions. More interestingly, the interviews supported the use of four interventions for which no evidence of effect on prescribing was found. These were 'peer review', 'local opinion leaders', 'practice support' and 'rules and obligations'. In that sense the qualitative study complemented the overview of systematic reviews. Two interventions with no evidence of effect on prescribing remained as such after the qualitative study: 'inter-professional education' and 'continuous quality improvement'.

The qualitative data were analysed again noting the higher level concepts of guideline implementation (Chapter 5). This analysis was used for the development of a simple model for implementing prescribing

recommendations in primary care. It comprised six steps: 'choose the condition', 'choose the guideline', 'identify influential people', 'identify organisational factors', 'plan and adopt the implementation strategy' and 'monitor the adherence'. The model provided a framework for planning the implementation of guidelines and recognising barriers that hindered adherence to guidelines.

The final phase of the project consisted of two surveys aiming to explore the ability of TPB to explain and predict GPs' adherence to guidelines. The qualitative study and the regional pilots were used to develop the questionnaires. Two stratified random samples of GPs in England were surveyed. Prescribing data was obtained from PACT. Data were collected from a total of 155 GPs for the statins survey. GPs responding to the statins survey were more likely to be working in practices with training status and probably less likely to be single-handed. Direct and indirect measures of attitude (behavioural beliefs) and indirect perceived controls (control beliefs), but not subjective norms (normative beliefs, peer pressure) were predictors of variation in prescribing intentions. TPB items explained up to 14% of variation in effective delivery outcome (Table 7.8). The main predictors were one behavioural belief item ('quality marker') and one self-identity item ('evidence-based practitioner'). Addition of practice and demographic variables increased the goodness-of-fit of the model to 39%. None of TPB measures were significant contributors to this latter model, however one control belief ('inclusion in the BNF') and one behavioural belief ('quality marker') approached significance. TPB items also explained up to 12% of variance in efficient statins prescribing indicator. A normative belief ('PCO advisers') and a control belief ('NSF advocates statins prescribing') were significant contributors to the model (Table 7.9).

122 GPs responded to the asthma surveys. The non-response analysis suggested that probably the only difference between responders and non-responders was in that the responders were more likely to be working in practices in which all partners graduated in the UK regardless of their ethnic origins. The variance in prescribing intentions was explained by subjective norms and perceived controls, and surprisingly attitudes had no contribution to the models. No TPB variable had any role in models for the effective delivery

primary outcome measure. On the other hand, a control belief ('BTS changed their recommendations') explained 9% of the variance in the secondary outcome (corticosteroid to bronchodilator ratio; Table 7.13). Three belief based items explained 22% of the variance in efficiency outcome (Table 7.14). They were 'patients will be healthier' (behavioural belief), 'hearing about asthma guidelines' (control belief) and 'PCO advisers' (normative belief).

The surveys generally supported the ability of the TPB for improving the understanding of GPs' behavioural, control and normative beliefs and their intentions. On the other hand, the surveys demonstrated the limits of the theory in identifying how different beliefs affected prescribing. The contributions of beliefs in observed variations in the prescribing outcomes were not mediated through the intentions, despite the TPB. This might have been caused by limitations of measuring intentions, measuring prescribing outcomes or both (7.4.4; 7.4.6 and 7.4.7). In both of the surveys perceived pressure from PCO advisers was correlated with less efficient prescribing (7.3.6 and 7.3.9). TPB advocates suggested that in studies with prospective designs (i.e. where behaviour was measured after beliefs and intentions), contribution of TPB variables in explaining variation of behaviour was causal (Ajzen, 1991; Conner and Sparks, 1996). Since the surveys employed prospective designs, then it might be concluded that PCO advisers caused less efficient prescribing. Despite Ajzen and others, the direction of cause and effect was not known. Although data was collected longitudinally, it did not prove cause and effect relationship (Grimes and Schulz, 2002). Alternatively, there might have been reverse causal relationship. It could be that GPs who prescribed statins and inhaled corticosteroids less efficiently were under more pressure from PCO advisers. The exact nature of relationship between these variables can be established in long term cohorts or interventional studies. The finding, however, was in line with the results of the overview and the qualitative study and questioned the effectiveness of educational outreach visit interventions.

The thesis also resulted in some methodological advances for the design of TPB surveys. Chapter 6 reported the findings of sample size calculations. This was the first known systematic attempt to determine sample size for TPB surveys. The sample size calculations were based on two different approaches:

the λ method (reported values of regression model goodness-of-fit) and the VIF method (variance inflation factor; zero-order correlations). The VIF was the suitable method. The methods used in Chapter 6 could be employed with adjustments in other settings and for other regression models.

8.3. Strengths and weaknesses of the study

Data from the overview of systematic reviews was analysed using narrative synthesis. Due care was given to ensure the analysis was performed objectively. Since only one investigator analysed the data, the possibility of bias in the analysis could not be ruled out. The overview was limited to published systematic reviews; therefore more recent evidence on behaviour change might have been missed. Also limitations in quality assessment approaches might have resulted in loss of understanding. The overview of systematic reviews had the advantage of not limiting itself to the text of systematic reviews, but also consulting primary studies where appropriate. This improved the understanding of the literature and increased the validity of the findings. First, several systematic reviews had included primary studies that did not answer the overview's questions or did not meet its inclusion criteria. So consulting primary studies was useful to ensure 'pears and apples' were not mixed. Second, many primary studies were included in more than one systematic review. Third, there were discrepancies in the reviews' conclusions and it was useful to identify the sources of the discrepancies.

Because of the focused approach, the overview included smaller number of systematic reviews than overviews published prior to it (Grimshaw et al, 2001). Focused approach might have resulted in omission of interventions with potentially useful implications for primary care prescribing. It was argued that the advantages of focused approach were further than the disadvantages, mainly as it prohibited the dilution of evidence owing to diverse behaviours and settings. Some previously published overviews of systematic reviews, including those focusing on primary care, did not follow systematic approaches (Conroy

and Shannon, 1995; Cantillon and Jones, 1999; Smith, 2000). Therefore the overview reported in this thesis was probably the first overview of systematic reviews in primary care that followed explicit methodology.

Purposive sampling was the biggest threat to the validity of the qualitative findings. It might have caused selection bias and limited the generalisability of the findings. On the other hand the inclusion of academics of primary care in the sample was advantageous in enhancing the depth of the study. The study was also limited in the fact that a single researcher analysed the data. The data also was not subjected to respondent validation. Framework analysis was a powerful analytical approach that helped exploit the collected data. The secondary analysis of the findings took this one step further and was used to suggest a simple model of guideline implementation for prescribing in primary care. The qualitative study was also different from previous studies in its focus on clinical guidelines for five clinical topics. It ensured variation in complexity of the disease, availability of clinical guidelines, GP role in prescribing and the importance of prescribing in management of the problem (4.2.3).

Despite the inherent subjectivity of methods used for the analysis of qualitative data, the validity of some of the qualitative study findings were established in studies published recently. For example the analysis suggested that computerised clinical guideline systems had limited use for the management of complicated clinical problems (e.g. hypertension or asthma) and were more appropriate for diseases with limited options (e.g. statins for CHD prevention, antibiotics for tonsillitis). Similarly a recent RCT concluded that computerised systems were not effective means of improving quality of care for management of asthma and angina in primary care (Eccles et al, 2002; Rousseau et al, 2003).

The main threats to the validity of the surveys were poor response rates. This was despite extensive attempts to boost the rates. Detailed non-response analysis revealed few differences between the responders and non-responders. Non-response analysis helped to understand the limits of generalisability of the findings. Despite the analyses one might argue that responders were different

from non-responders in the fact they were willing to contribute a few minutes of their time to a research study. 'Altruistic' tendencies of GPs might be correlated with the service they provided (Le Grand, 2003). Non-response is a challenge to studies in primary care, especially TPB studies as they require long questionnaires. Another limitation of the surveys was due to the limited validity of prescribing outcomes. The validity of PACT data in terms of the appropriateness and adequacy of prescribing for individual patients was limited. Prescribing units captured some of the variation caused by case-mix differences, but this was done at practice level and not for individual patients. This could have been improved by accessing data that linked patients with prescribing. Efficiency indicators were more valid than effective delivery indicators of prescribing. Interestingly, TPB items had more explanatory powers in models of prescribing efficiency. The regional pilots helped improving the surveys' tools and contributed to the objective choice of optimal scaling approaches for the analyses. The uncertainties in optimal scaling of TPB items were hardly addressed in previous studies of health care providers. Explicit sample size calculations added to the available literature on TPB.

Previous TPB studies of health professionals did not followed the methods used in the surveys. The surveys were designed prospectively and were not limited to measuring behavioural intentions (e.g. Walker et al, 2001; Puffer and Rashidian, 2004) or self-reported behaviours only (e.g. Millstein, 1996; Levin, 1999; Watson and Myers, 2001). The surveys were the first known TPB studies of GP adherence to clinical guidelines. Prior to the surveys, there was only one TPB study of GP prescribing, focusing on intentions and not prescribing (Walker et al, 2001). Another study tested the theory of reasoned action in explaining variation in physician prescribing, but was grossly underpowered (Lambert et al, 1997). Both studies were concerned with antibiotic prescribing. Walker et al (2001) achieved higher response rate than the asthma and statins surveys. Its investigators chose not to include some TPB elements in their questionnaire (i.e. direct subjective norms). Also the major difference between that study and the surveys was that Walker et al did not request any access to GP prescribing data.

The development of the questionnaires was not based on a brief set of interviews, a common approach in TPB studies. Rather the questions were developed using the qualitative study. The framework analysis of qualitative data also developed themes for clinical guideline implementation in primary care and some were supported by the TPB surveys. For example the qualitative study suggested that change in clinical guidelines might hinder implementation of evidence base recommendations. The analysis of the secondary asthma prescribing outcome concluded that a control belief that 'BTS asthma guidelines had substantially changed their recommendations' was significantly correlated with less effective prescribing. The qualitative analysis also supported the TPB assertion that beliefs had important roles in the formation of intentions and behaviours.

The choice of clinical conditions for the qualitative study and in particular for the surveys was deliberate and objective. Although the choices were made for technical considerations, CHD prevention and management of asthma were important public health issues with high burden of disease and cost to the society and individuals. The findings of the study could contribute to better care for imperative conditions with potentials for reducing unnecessary costs.

8.4. What does it mean? The study implications

8.4.1. The implications for future research

As noted before, optimal scaling methods for the TPB studies were not established. Analyses performed for the surveys should contribute to the establishment of optimal scaling approaches for the theory. Future studies might consider statistical optimisation approaches (Box and Cox, 1964). They should consider uncertainties on the natural position of zero on the scales, and ideally provide room for variation of the zero point on both expectancy and value items of belief based measures (French and Hankins, 2003).

The TPB studies were helpful in identifying important beliefs and the values assigned to those beliefs. TPB tools and questionnaires however were not sensitive enough to identify delicate differences between groups of responders. GPs might approach clinical guideline recommendations selectively. Hence, it is useful to know which guideline recommendations are implemented by what group of practitioners. Then it will be useful to know the characteristics (e.g. beliefs) of GPs that implement or do not implement certain recommendations. In marketing research, this separation of the target population into subgroups is known as 'segmentation' (Maibach et al, 2002). Sophisticated choice modelling ('conjoint analysis') methods are developed to identify and understand the existence of segments and subgroups (Ryan, 1999; Ryan and Farrar, 2000; Gustafsson et al, 2003). Identification of the subgroups can help devising different interventions specific to the needs of different subgroups. Choice modelling may use discrete choice, rating or scaling questions and may or may not incorporate vignettes (case scenarios). Choice modelling approaches are previously used in TPB studies but mainly as a measurement approach for the expectancy-value elements of the theory (Jonas, 1993). Choice modelling is also recommended as a technical solution to the problem of scaling the belief based items (French and Hankins, 2003). The suggestion proposed here differs from previous studies as it aims to use choice

modelling for the identification of subgroups within the target population and not just as technical remedy to statistical problems.

Despite limitations, significant proportions of prescribing outcomes were explained by TPB variables. This can be improved by objective measurement of prescribing outcomes in future studies. Difficulties in the measurement of volitional prescribing by GPs suggest that future research should test the theory by measuring prescribing for new diagnoses or prescriptions and should exclude repeat prescriptions (7.4.4). Conducting TPB studies alongside RCTs can improve the quality of data for testing TPB. It will also help identifying reasons for adopting or not adopting interventions.

Choice modelling might help enhancing the findings of the thesis in other ways. The qualitative study and the asthma survey suggested that change in guideline recommendations might negatively affect implementation. On the other hand, changes are inevitable. It will be useful to quantify the adverse effects of changes in recommendations on adherence rate, as it can inform guideline development as well as guideline implementation processes. Quantification of change effects on implementation does require experimental research designs, for example by changing guidelines' recommendations in different ways, implementing them in different geographical areas and then evaluating the implementation. These experimental studies will be very difficult in practice and may face political, professional and ethical resistance. Choice modelling can be used to assess the effects of hypothetical clinical guideline recommendations on stated implementation intentions.

Efforts should be made to increase the understanding of the validity of behavioural theories using empirical research. Throughout the period that this thesis was undertaken, the Medical Research Council funded an investigatory study assessing the validity of a few psychological theories for provider behaviour change. Empirical investigations should be expanded to the theories developed by other disciplines (Chapter 3).

It seems that the effect sizes of behaviour changes interventions are declining. Although previous behaviour change improvement studies were likely

to achieve significant improvements (North of England Study of Standards and Performance in General Practice, 1992a; North of England Study of Standards and Performance in General Practice, 1992b), more recent trials tend to achieve limited success (Freemantle et al, 2002; Wright et al, 2003). Theory-based approaches can help understanding why. TPB can be used for devising behaviour change interventions (Ajzen, 2002a), yet the effectiveness of TPB interventions for changing provider behaviour is not known (Hardeman et al, 2002). A postal TPB intervention successfully changed molar tooth extraction *intentions* in the intervention group of dentists (Bonetti et al, 2003). The findings of the asthma and statins surveys and of this trial indicate that TPB has potential for contribution to active implementation of provider behaviour change strategies. This needs to be investigated in carefully designed RCTs. For that reason, a protocol for a randomised trial assessing the potentials of TPB based educational intervention was prepared (Appendix VIII). As well as the effectiveness trial, the proposed study included a cost-effectiveness analysis (Drummond et al, 1997; Coyle et al, 1998). This would be important, in particular if the potential effect sizes were small. Even small changes in prescribing might have important effects on health care costs as well as patient outcomes (Jamtvedt et al, 2003).

8.4.2. The implications for guideline implementation in middle-income countries

Before discussing the implications of the findings for middle-income countries, a few issues should be noted. The nature of prescribing problems in developing countries is somehow different from what is usually observed in the UK. In many countries poly-pharmacy or multi-drug use is a challenge to the system. Studies have shown that the average number of drugs per prescription is between 2.4 to ten, while if prescribing is conducted rationally, the average number of drugs per prescription should be less than two (Le Grand et al, 1999, Dinarvand and Nikzad, 2000, Hafeez et al, 2004). A random survey study of 55 pharmacies in Tehran observed that about 40% of patients who were served by daily pharmacies and 80% of those served by 24-hour service pharmacies were

not aware of the recommended dosages of the prescribed items. This was complicated further by the fact that the average time the pharmacists spent with each patient was only 35 seconds (Dinarvand and Nikzad, 2000). Another frequently cited prescribing challenge is the overuse of injections. For example studies demonstrated that 15-60% of people attending health care facilities in different developing countries received at least one injection (Hadiyono et al, 1996, Dinarvand and Nikzad, 2000, Hafeez et al, 2004). Also it was reported that 40-50% of prescriptions contained antibiotics (Dinarvand and Nikzad, 2000; Hafeez et al, 2004). The private sector is a major provider in many of these countries, especially of primary care, and is very much unregulated. The scarce evidence available from developing countries is usually from the public sector (Le Grand et al, 1999). All these factors suggest that clinical guideline implementation and quality improvement in developing countries should probably have a different focus than organisations such as NICE.

The results of the studies of prescribing in the UK, therefore, are not easily generalised to middle-income developing countries. The main implications of the qualitative study and the surveys are that they provide a model for evaluation that can be applied elsewhere. Although the focus of behaviour change may be different in developing countries – for example use of statins may be a future concern – it is likely that the systems will face problems similar to those reported in the qualitative study if they target use of e.g. aspirin for CHD prevention or effective management of diabetes and hypertension. The thematic framework and qualitative model can be useful, as they illustrate the importance of different factors that may affect guideline implementation. TPB is likely to be beneficial in explaining variation in the uptake of guidelines or in prescribing.

The results of the overview of systematic reviews reported positive findings from interventions conducted in developing countries. The included studies reported evidence from developing countries that suggested the trials of interactive educational interventions and educational outreach visits achieved significant improvements in provider behaviour (2.3.3 and 2.3.8). In order to understand the generalisability of the overview findings, other common characteristics of health systems in developing countries should be noted. The

cost effectiveness of the interventions to change physician behaviour in low and middle-income countries can be different from high-income countries. For example in high-income countries an intervention like educational outreach is costly, because of high personnel costs (in comparison to other services and materials). As many high-income countries have a shortage of qualified nurses and doctors, opportunity costs are even higher. In comparison, staff cost in developing countries are relatively lower than drugs or instruments. Therefore interventions that require a lot of staff time may turn out cheaper than those that require other services and material. For example an RCT of improving prescribing for diarrhoea in Indonesia found educational outreach were much cheaper than seminars; i.e. less than a quarter ($0.77 versus $3.30 per participant; Santoso et al, 1996). In some middle-income developing countries, there are more qualified health professionals than the system requires (Malekzadeh et al, 2001; Shadpour, 2003). Therefore investment of resources on effective interventions that require active involvement of health care providers (i.e. labour intensive interventions) is potentially rewarding.

8.4.3. The implications for guideline implementation in the UK

Several findings of the thesis had direct implications for guideline implementation in the UK.

Interventions based on management theories (e.g. continuous quality improvement and practice support) had weak evidence base (Table 2.1). This might be due to the discipline's culture (e.g. tendency for not using experimental designs) or due to the difficulties of subjecting these interventions to experimental designs. Practice support, e.g. by developing mini-clinics, was perceived effective in the qualitative study, but not in the surveys. The surveys provided mixed messages. They found GPs who reported having CHD clinics prescribed statins more efficiently. On the other hand, GPs who reported having asthma clinics provided less efficient prescribing and potentially of lower quality. The caveats of these findings were discussed in Chapter 7. However, the clear

message is that rigorous assessment of the effectiveness of interventions based on management theories is warranted.

Interactive continuous medical education should be encouraged and replace didactic CME. There was evidence of effect for the former, and evidence of no effect for the latter. The only caveat in here is that the evidence of effect for interactive CME was partly originated from developing countries, hence it might be setting sensitive. It is also likely to be costly, especially in the UK, so its cost-effectiveness should be assessed. CME programmes were barely mentioned in the interviews. The overview also pointed out that interventions based on financial incentives were likely to have unwanted outcomes. Although selective use of financial incentives might be advocated, as noted in the qualitative study, its blanket use should be avoided (Rashidian et al, 2005).

Do the findings discourage the usage of educational outreach visits? This question is difficult to answer. Throughout the period when this thesis was undertaken, evidence about the effectiveness of educational outreach visits has been growing. It is no longer possible to say that educational outreach visits are effective in changing prescribing, since more recent trials provide a mixed picture. Moreover, the qualitative study participants had mixed feelings towards the role of prescribing advisers in primary care. Efficiency indicators' analyses for the asthma and statins surveys suggested GPs who perceived more strongly that PCO prescribing advisers expected them to follow clinical guidelines, were more likely to prescribe less efficiently. This did not provide evidence of causal effect (see 8.2); nonetheless, the findings of the surveys as well as the qualitative study and the overview of systematic reviews complemented each other in highlighting uncertainties. Therefore, in answer to the question, one might argue that policy-makers should proceed with caution. They should limit the amount of resources spent on different forms of educational outreach visits; use the intervention for selected areas of care depending on local and national priorities, and further evaluate the intervention.

The findings of the qualitative analyses provided simple criteria for identification of conditions for which prescribing guidelines might not be

warranted ('choose the condition'; 5.3.1 and Fig 5.1). With the current level of enthusiasm for guideline development, such findings should be helpful. The qualitative analysis also noted the importance of 'first contact' (i.e. if GPs perceived a guideline as useless, it would more difficult to gain their confidence back; 4.3.9). Hence, the qualitative study supported the care that NICE gives to refining guidelines' recommendations before publication. In NICE, this process is called 'consultation' phase. Obviously, it should be noted that consultations might also negatively result in deviating recommendations away from evidence; however, they are great opportunities to ensure guidelines would not fail in 'first contact'.

The qualitative study indicated that inclusion of guidelines in the BNF improved implementation. Also the NSFs' recommendations were likely to affect prescribing. Both these findings were supported in the statins survey. BNF is potentially an under-used source for guideline implementation. It is therefore recommended that BNF should refer to more guidelines when discussing prescribing alternatives, especially to guidelines for the care of chronic diseases.

The notion of change in clinical guideline recommendations was highlighted both in the qualitative study and the asthma survey. Some interviewees did not welcome change in guidelines recommendations, in particular if they were not convinced why change in recommendations was required (4.3.3). Similarly, the GPs who agreed more strongly with the statement: 'BTS asthma guidelines had substantially changed their recommendations', had potentially lower quality of prescribing for asthma (7.3.9). Given evidence-based clinical guidelines are expected to be updated in pre-specified intervals (Shekelle et al, 2001), further investigation of the effects of change in recommendations on implementation is warranted. The immediate implication, however, is that updating guidelines should not be seen as mere updating of evidence-based recommendations. Rather, guideline developers should clearly justify the changes made in their recommendations and effectively communicate the change with providers.

The very first section of the chapter revisited the complexity of changing prescribing behaviour. Despite the complexities, all interested parties want to see improvements in quality of care. It is sad that evidence based messages are not implemented despite all the good wills and efforts. One of the interviewees said: '*we as professionals want to know we are doing our best for our patients*'. Can this thesis be of any help? Emerson once said '*sometimes a scream is better than a thesis*' (Persaud, 2000). Patients may use occasional 'screams' to get the care they expect. The thesis may not eliminate the need for occasional screams – by patients or others – but might contribute to reducing the need for it. Simple interventions can be cost-effective ways of implementing guidelines and should not be ruled-out. GPs views and beliefs should not be ignored. Understanding GPs' beliefs may improve the understanding of their prescribing and enhance the effectiveness of interventions to improve care.

Appendices

*Appendix II-1. Search strategy for identifying systematic reviews of provider behaviour change in Medline database**

1 exp *education,continuing/

2 (education$ adj2 (program$ or intervention? or meeting? or session? or strateg$ or workshop? or visit?)).tw.

3 *pamphlets/

4 (behavio?r$ adj2 intervention?).tw.

5 (leaflet? or booklet? or poster or posters).tw.

6 ((written or printed or oral) adj information).tw.

7 (information$ adj2 campaign?).tw.

8 (education$ adj1 (method? or material?)).tw.

9 outreach.tw.

10 ((opinion or education$ or influential) adj1 leader?).tw.

11 facilitator?.tw.

12 academic detailing.tw.

13 consensus conference?.tw.

14 practice guideline?.tw.

15 (guideline? adj2 (introduc$ or issu$ or impact or effect? or disseminat$ or distribut$)).tw.

16 ((effect? or impact or evaluat$ or introduc$ or compar$) adj2 training program$).tw.

17 *reminder systems/

18 reminder?.tw.

19 (recall adj2 system$).tw.

20 (prompter? or prompting).tw.

21 algorith?.tw.

22 *feedback/ or feedback.tw.

23 (feedback adj1 (loop? or control? or regula$ or mechanism? or inhib$ or system? or circuit? or sensory or visual or audit$)).tw.

24 22 not 23

25 chart review$.tw.

26 ((effect? or impact or records or chart?) adj2 audit).tw.

27 *patient education/

28 marketing.tw.

29 ((effect? or impact or evaluat? or introduc$ or compara$) adj2 (prevent$ program$ or screening program$)).tw.

30 ((introduc$ or impact or effect? or implement$ or computer$) adj2 protocol?).tw.

31 (computer$ adj2 (dosage or dosing or diagnosis or therapy or decision?)).tw.

32 *physician s practice patterns/

33 (meta-analysis or review literature).sh.

34 meta-analy$.tw.

35 metaanal$.tw.

36 (systematic$ adj4 (review$ or overview$)).tw.

37 meta-analysis.pt.

38 review.pt.

39 case report/

40 letter.pt.

41 historical article.pt.

42 review of reported cases.pt.

43 review,multicase.pt.

44 review.ti.

45 review literature.pt.

46 33 or 34 or 35 or 36 or 37 or 38 or 44 or 45

47 39 or 40 or 41 or 42 or 43

48 46 not 47

49 animal/

50 human/

51 49 not (49 and 50)

52 48 not 51

53 or/1-21,24-32

54 52 and 53

* Source: Grimshaw et al, 2001.

Appendix II-2. Evidence table for the overview of systematic reviews

(descending in time period order)*

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Jamtvedt et al, 2003)	**Study designs** RCT **Time period** 1966 – 2001 (last substantial amendment in 2003) **Participants** Health professionals **Intervention** Audit and feedback vs no interventions and other interventions **Main outcomes** Provider performance in health care setting and patient outcome. Excluding performance in test situation.	85 papers included. 18 studies were on prescribing. 57 studies in primary care. In an American study of 33 PCPs audit and feedback plus educational outreach significantly reduced inappropriate tetracycline prescribing for upper respiratory infection. Control group received none (McConnell et al, 1982). A multi-centre study of compliance with guidelines for asthma in four countries found small positive effects in three countries and no effect in one (Veninga et al, 1999). The intervention groups received audit and feedback plus educational meeting versus no intervention. Two studies that compared audit and feedback as part of a multi-faceted intervention with no intervention observed no effect. In one study the main outcome measure was the average number of drugs prescribed in which a national treatment guideline was introduced in Uganda (Kafuko et al, 1999). The second study of 20 general practices focused on compliance with guidelines for high cholesterol (van der Weijden et al, 1999). [This latter study has no prescribing outcome] One study with four arms compared audit and feedback alone to audit and feedback plus patient education materials to improve antibiotic prescribing for respiratory infections in which adding patient education had no effect In another comparison audit and feedback was compared to patient education and there was no difference between groups in antibiotic prescribing (Mainous et al, 2000). A three arm study compared a single letter recommending that the number of medications received by patients should be reduced with audit and feedback plus a compliance index, peer review and recommendations; and to no intervention. At four months both intervention groups had significant reductions in poly-pharmacy compared to the control, but there was no difference between the two interventions (Meyer et al, 1991). [this study is performed in hospital outpatient clinic, included nurses]	No general conclusion specifically relevant to prescribing. The review suggests that the audit and feedback may be effective, but the evidence to support its effectiveness is very patchy. The effects are generally small to moderate. Of the variables that the reviewers specified, baseline compliance was the only one that helped to explain variation in the absolute effectiveness across studies and possibly, to a lesser extent, in the relative effectiveness. In contrast to the conclusions of other systematic reviews the reviewers did not find evidence of a larger effect for multi-faceted interventions compared with audit and feedback alone. One study reported a large effect of audit and feedback plus outreach (McConnell et al, 1982). This does not suggest that audit and feedback	In order to collate the findings of different studies, the reviewers categorised all studies in terms of the complexity of the target behaviour. They categorised behaviours such as antibiotic prescribing in low complexity group, and guideline implementation in moderate complexity group. Categorisation of all guideline implementation studies in the moderate complexity group seems simplistic and may explain why no relationship between the complexity of the target behaviour and effectiveness was observed.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
		One study that compared feedback on medication with feedback on performance found no difference in control of blood pressure (Gullion 1988). [check for prescribing]	alone or in combination with educational meetings is likely to have large effects.	
		A high quality study that compared audit and feedback with no intervention found audit and feedback had no impact on GPs prescribing (O'Connell et al, 1999).		
(Gross and Pujat, 2001)	**Study designs** Retrospective and prospective. Controlled and uncontrolled interventions **Time period** 1966 – Dec 2000 **Participants** Prescribers in primary, secondary or tertiary care **Intervention** Implementation of clinical guidelines for improving antibiotic use for common infectious diseases, HIV/AIDS, and vancomycin prescription **Main outcomes** Antibiotic prescribing.	40 studies included. Eight studies meet our inclusion criteria. Other studies were excluded because of inappropriate designs, or non primary care settings. Antibiotics for upper respiratory tract infection: two CBAs of multi-faceted interventions observed improvement and an RCT using academic detailing found no improvement. An Australian RCT of 285 GPs found not much extra improvement in the intervention group which received academic detailing and educational materials for treatment of tonsillitis over improvement observed in control group (De Santis et al, 1994). A CBA trial of a multi-faceted intervention in 18 Mexico City primary care facilities (119 physicians) studied antibiotic prescribing for acute rhinopharyngitis (Perez-Cuevas et al, 1996). They observed reduction in the use of antibiotics and increase in the appropriate use of antibiotics in the intervention group. No change was observed in the control group. Another CBA trial of 56 physicians (and non physicians) in four sites studied antibiotic prescribing for acute bronchitis (Gonzales et al, 1999). The authors observed reduction in the use of antibiotics in the intervention site which received a multi-facet intervention including academic detailing. There was no increase in prescription of antibiotics for other diseases or increase in the re-visit rate of patients with bronchitis in this site. A multi-faceted CBA study resulted in increased appropriate antibiotic use for UTI while decreasing urinalysis and urine culture tests, all in line with	Multi-faceted interventions are more effective. Academic detailing (educational outreach), feedback (from nurses, pharmacists, physicians), local adaptation of a guideline, small-group interactive sessions, and computer-assisted care seems to be effective. Educational material mailouts seem to be ineffective.	The review does not note the level of analysis error in the primary studies. Some included multi-arm interventional studies report no significant improvement over control group in the mail out only group. This may be due to type II error. One large RCT found mailed out feedback and educational material an effective way of improving prescribing. At least three eligible studies of changing antibiotic prescribing included in other (previous) reviews were not included in here (Klein et al, 1981; McConnell et al, 1982; Angunawela et al, 1991).

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
		the guideline's recommendations (Saint et al, 1999). An American RCT of 435 prescribers focusing on three drugs found academic detailing a cost-effective method of improving prescribing and reducing cephalosporine prescriptions when compared with educational materials mailout group and control group (Avorn and Soumerai, 1983). They found reduction in mailout group in comparison to control, but the change was non significant. A further CBA in the USA targeted antibiotic prescribing and offered pharmacist-led academic detailing to physicians in one region, physician-led academic detailing to another region, educational materials mailout to the third region, and the fourth region was used as control (Schaffner et al, 1983). They observed reduction in prescriptions in all four regions. Reduction in prescribing in both academic detailing groups was significantly lower than the control group, and more so for the physician-led group. A third study also focused on pharmacist-led academic detailing in an RCT of 112 Australian GPs (Ilett et al, 2000). They focused on upper and lower respiratory tract infections and UTIs. They observed significant increase in prescribing for amoxicillin and doxycycline (cheap and effective drugs), and decrease in prescribing for cefaclor and roxithromycin (more expensive antibiotics) in the intervention group. A Canadian RCT of 1095 primary care providers mailed out individualised feedbacks to each provider in the intervention group as well as educational materials twice (Hux et al, 1999). The feedback focused on providers prescribing for patients older than 65. The control group received none. The study observed increase in first line drugs use, and decrease in prescribing costs in the intervention group.		
(Grilli et al, 2002)	**Study designs** RCT, CBA, CCT, ITS **Setting** All settings **Time period** 1966 – 1999. Last substantive amendment in 2001.	Twenty studies were included. All used ITS designs. All the studies reported significant improvements, but not all were supported in the re-analysis performed by the reviewers. Only two studies were on prescribing: This Canadian study reported significant reduction in the use of calcium-channel blockers as the first line of treatment for hypertension. The finding was confirmed by re-analysis performed by the reviewers (Maclure et al, 1998). Another ITS study used mass media campaign to reduce use of aspirin in	Media campaigns may have a positive influence upon the manner in which health services are utilised. The findings support the importance of efforts to ensure that reporting of health-related issues in the lay media correctly represents the best available knowledge on the effectiveness of health care interventions.	Very few studies with acceptable quality have attempted to evaluate the effects of mass media campaigning on prescribing. Both included studies of prescribing showed significant improvement in the target behaviour. Although most targeted areas other than prescribing were meant

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	Participants Health care professionals, patients, public **Intervention** Mass media interventions **Main outcomes** Utilisation of health services	children to prevent Reye's syndrome. The study observed significant reduction in the use of aspirin, which was confirmed in re-analysis by the authors (Soumerai et al, 1992).		to increase health care utilisation (and this has been reflected in the review's synopsis), both prescribing campaigns were planned to reduce inappropriate prescribing and replace it with appropriate medicines.
(Lewin et al, 2001)	**Study designs** RCT, CCT, BCA, ITS **Setting** All health care settings **Time period** 1966 – 1999 (last substantive amendment in 2001) **Participants** Health care providers, including in-training **Intervention** Any interventions to promote patient-centred health care **Main outcomes** Consultation process, health care behaviour, patient outcome	17 studies were included. Only one study reported prescribing data in primary care: In an RCT, patient-centred training for providers plus condition-specific training for both providers and patients was evaluated against no intervention. Parents of children in the intervention were significantly more likely to report that the paediatrician had prescribed inhaled anti-inflammatory medicine for the child (82.7% vs 70.3%, p< 0.018) and also that the paediatrician had given the family a written plan for adjusting the doses of the medicine at home when symptoms change (26% vs 16%, p<0.03) (Clark et al, 1998).	There is some evidence that training health care providers in patient-centred approaches may improve satisfaction with care. It is difficult to quantify the effects of training health care providers in patient-centred care on health status.	

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Zwarenstein and Bryant, 2000)	**Study designs** RCT, CBA, ITS **Setting** Primary care, hospitals **Time period** 1966 – 1999 (last substantive amendment in 2000) **Participants** Doctors, nurses **Intervention** Interventions to improve doctor-nurse relationship **Main outcomes** Process outcomes, patient outcomes	Two studies were included, both from hospital setting. No study meets the overview's inclusion criteria.	Increasing collaboration improved outcomes of importance to patients and to health care managers. These gains were moderate and affected health care processes rather than outcomes. Further research is needed to confirm these findings.	Lack of any evidence may be due to the complexity and difficulty of devising interventions that can be implemented in practice.
(Armour et al, 2001)	**Study designs** RCT, ITS, observational **Setting** All setting **Time period** Not stated. No paper later than 1999 is included. **Participants** Physicians	Seven studies were included, of which two were RCTs and one ITS. None was on prescribing in primary care.	Explicit financial incentives that place individual physicians at financial risk can be effective in influencing physician resource use. However, the empirical evidence regarding the effectiveness of bonus payments on physician resource use is mixed. Similarly, our review disclosed mixed effects of the influence of explicit financial incentives on the quality of patient medical care.	The authors state in the paper they have followed the Cochrane Handbook methods and approaches for the review. The evidence base of the conclusions is not very clear since some of the included studies were of weak designs. None of the conclusions applies to the overview because of the lack of relevant evidence.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	Intervention Financial incentives **Main outcomes** Resource utilisation, provider performance			
(Chaix-Couturier et al, 2000)	**Study designs** Almost all types of studies (experimental and non experimental; prospective and non prospective) **Setting** All settings **Time period** Jan 1993 - May 1999 **Participants** Physicians **Intervention** Financial incentives **Main outcomes** Provider performance; patient outcomes	89 studies included, only eight were RCTs. An RCT of the effects of disclosure of thresholds for financial sanctions to providers (mainly physicians) concluded that disclosure of the threshold resulted in reduction in prescriptions (Nyman et al, 1995).	High quality evidence is lacking to establish the effectiveness of different remuneration systems in changing provider behaviour. Financial incentives should be used carefully as they may cause conflict of interest between different providers and also between different sections of health system and public. "Any form of fund-holding or capitation decreased the total volume of prescriptions by 0-24% ... compared with fee-for-service."	The review's main conclusion of the effects of capitation and fund-holding on prescribing is originated from low quality studies (this is acknowledged by the authors).
(Ratanawijitrasin et al, 2001)	**Study designs** All research designs **Time period** 1966 – 1999	18 studies were included. Almost all were post-intervention studies. Only two were before-after designs, both without control groups. None of the included studies meet the overview's inclusion criteria.	None of the results are conclusive because of weak designs. Findings suggest an association between increases in the supply of essential drugs (plus training) and more appropriate use of	The review fails to identify any evidence of effectiveness (or ineffectiveness) national drug policies in developing countries. Carefully written and good quality study.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	Participants Governments, policy-makers, providers **Intervention** Laws, rules, administrative orders, national drug policies, and essential drug programmes in developing countries **Main outcomes** Provider behaviour, health systems objectives		medications in primary care. It suggests some unintended effects of de-registration of drugs or upward reclassification of specific medicines. Similarly, loosening restrictions have sometimes been accompanied by increased dispensing of specific drugs by unqualified personnel.	
(Thomson O'Brien et al, 2001)	**Study designs** RCT, CBA **Setting** All settings **Time period** 1966 – Jan 1999. Last substantive amendment in 2000 **Participants** Health care professionals, and those in post-graduate training **Intervention** Continuing educational meetings and workshops **Main outcomes**	32 papers (30 RCTs) coveting 36 comparisons were included. Four included papers focus on prescribing in primary care. A study of using didactic approach in Sri Lanka had small non-significant effects on injection rates.(Angunawela et al, 1991). Workshop used in Indonesia was effective with moderate to large effects on intramuscular injection rate in Indonesia. The results were statistically significant (Hadiyono et al, 1996) Mixed methods used in Zambia to improve prescribing, with small but statistically significant effects (Bexell et al, 1996). An RCT of a theory based educational seminar aimed at general paediatricians in community practices improved asthmatic patients' outcomes and practitioners' prescribing (Clark et al, 1998) (after checking the original paper).	Interactive workshops can result in moderately large changes in professional practice. Didactic sessions alone are unlikely to change professional practice.	This is an update of a series of previous reviews (Davis et al, 1992; Davis et al, 1995; Davis et al, 1999). Older studies have a wider inclusion criteria and include interventions other than CME (Davis et al, 1992; Davis et al, 1995). The paper does not discuss the findings of individual papers, probably because of the large number of included studies. Instead they are discussed as a whole in the findings section. Therefore it is difficult top extract messages relevant to prescribing in primary care settings. At least four included papers focus on prescribing. The message derived from these studies is in line with the overall conclusions of the reviewers.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	Provider performance; patient outcome			
(Zwarenstein et al, 2000) **Journal publication: (Zwarenstein et al, 1999)**	**Study designs** RCT, CBA, ITS **Setting** All health care settings **Time period** 1962 -1998 (last substantive amendment in 2000) **Participants** Health and social care professionals **Intervention** Inter-professional education **Main outcomes** Health care process measures, patient (or vignette) outcomes (reported or measured)	No study met the inclusion criteria		Lack of any evidence may be due to the complexity and difficulty of devising interventions that can be implemented in practice.
(Bower and Sibbald, 1999) **(Bower and Sibbald, 2000)**	**Study designs** RCT, BCA, ITS **Setting** Primary care **Time period**	38 studies were included. All were conducted in primary care setting. 28 included studies meet the overview's inclusion criteria, and are mostly from the UK. **Studies of 'direct' effects** Replacement models: Prescribing outcomes were grouped in terms of whether they dealt with	General: This review does not support the hypothesis that adding mental health workers to primary care in 'replacement' models causes a significant or enduring change in	Two of the included studies seems to be duplicate publications (Katon et al, 1995; Katon et al, 1996). At least they seem two reports of one intervention. Similarly for two studies only one reference being provided and it seems that one of them is a more complete reporting of the first one

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	1966 – 1998 (last substantive amendment in 1999) **Participants** PCP, mental health worker, other primary care providers **Intervention** On-site mental health workers in primary care (substitution or consultation-liaison models) **Main outcomes** Referral, prescribing, diagnosis	short term (six months or less) or long term (more than six months). Short term effects on psychotropic prescribing rates or costs were examined in eleven studies. Four studies (Earll and Kincey, 1982; Robson et al, 1984; Boot et al, 1994; Mynors-Wallis et al, 1997) showed significant reductions in the intervention group. Two studies found no significant difference (Corney, 1984; Hemmings, 1997). Six studies that did not report statistical significance. Four studies examined the short term effects on non-psychotropic prescribing. Two studies reported no significant differences (Earll and Kincey, 1982; Robson et al, 1984). Two studies did not report significance. Long term effects on psychotropic prescribing rates were examined in six studies, three of which found no significant differences (Earll and Kincey, 1982; Robson et al, 1984; Hemmings, 1997). One study found a significant difference in favour of the intervention group in terms of tranquillisers but not antidepressants (Ashurst and Ward, 1983). Two studies did not report significance. Four studies examined the long term effect on non-psychotropic prescriptions. Three reported no significant differences between groups (Earll and Kincey, 1982; Robson et al, 1984; Ginsberg et al, 1984), one did not test for significance. Consultation-liaison models Six studies examined effects on short term medication use. Three studies (from one group) found that intervention patients were more likely to receive an adequate duration or dosage of drugs (Katon et al, 1992a; Katon et al, 1995; Katon et al, 1996), although in one the change was observed in only a sub-group of patients (Katon et al, 1996). Three other studies found no significant differences in rates of prescription, patient adherence, or dosage received (Wilkinson et al, 1993; Mann et al, 1998). Two studies examined effects on long term medication use. One found no significant difference in the proportion of patients who filled a prescription, but intervention patients were more likely to have filled three or more prescriptions (Katon et al, 1992a). The second study found that intervention patients were more likely to receive a change in medication and to have higher prescribing costs (Katon et al, 1995). **Studies of 'indirect' effects**	PCP behaviour. 'Consultation-liaison' interventions may cause changes in psychotropic prescribing, but these seem short-term and limited to patients under the direct care of the mental health worker. Longer-term studies are needed to assess the degree to which demonstrated effects endure over time. Replacement models In terms of the 'direct' effects of 'replacement' models, evidence concerning prescribing behaviour is also inconsistent, but suggests that reduction in psychotropic prescription may be achieved in the short term: there is less evidence to support long term effects. In terms of the 'indirect' effects of 'replacement' models, the results of a small number of studies are consistent in showing that there is no reduction in psychotropic prescribing. Consultation-liaison models	(Mann et al, 1998).

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
		Replacement models: Three CBA studies examined the 'indirect' effects on prescribing and found no effect because of the intervention (Pharoah, 1996; Coe et al, 1996; Baker et al, 1996). Consultation-liaison models Two CBA studies examined the indirect effect on prescribing behaviour. In the first, PCPs significantly increased their rates of antidepressant prescription post-intervention, whereas control clinic did not (Katon et al, 1992b). In the second study no clinically important difference was observed.	Direct effects: changes in antidepressant prescribing in both the short and long term are not consistent across studies, but, where outcomes are significant, this is related to multi-faceted consultation-liaison interventions which involve education, patient-based consultation and feedback and reorganisation of services. Indirect effects: There is only weak evidence that 'consultation-liaison' models may affect the rates or type of antidepressant prescribing in the wider patient population. There is no evidence of any influence on the adequacy of medication, or that effects persist past the end of the intervention.	
(Hunt et al, 1998)	**Study designs** RCT, CCT, CBA **Setting** All settings **Time period** 1974 – 1998 **Participants** Physicians and other health care professionals, in practice or training	68 studies were included. Most happened in hospital settings. Nine out of 15 studies of computer assisted dosing of toxic drugs were presented different degrees of effectiveness, but all were conducted in hospital settings. Two studies conducted in outpatient clinics (still not primary care) were ineffective. CDSS was more effective for intravenous administered toxic agents, and less for oral agents (warfarin).	There is strong evidence that CDSSs can improve quality of care offered to patients. CDSSs are most effective as reminder systems for preventive care. No improvement was observed in studies evaluating drug prescribing.	It is an update of (Johnston et al, 1994). The authors' main conclusion does not apply to prescribing in primary care since virtually no primary study of prescribing in primary care was included in the review. However, the effectiveness of the interventions for hospital prescribing was not established in the review.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	Intervention Computer-based clinical decision support systems or 'expert systems' **Main outcomes** Clinician performance; patient outcome			
(Thomson O'Brien et al, 1999)	**Study designs** RCT **Setting** All settings **Time period** 1966 – May 1998 **Participants** Health care professionals (excluding students only studies) **Intervention** Local opinion leaders **Main outcomes** Provider performance; health care outcome (excluding studies that measure knowledge, or test performance only)	Eight trials were included. In seven trials local opinion leaders were identified by peers using a sociometric questionnaire in which humanity, communication and knowledge are important attributes (Hiss et al, 1978). All studies were conducted in hospital settings (some in community hospitals).	Using local opinion leaders results in mixed effects on professional practice. However, it is not always clear what local opinion leaders do and replicable descriptions are needed. Further research is required to determine if opinion leaders can be identified and in which circumstances they are likely to influence the practice of their peers. Only two trials show clinically important effects owing to local opinion leaders.	The review also states that finding opinion leaders may be difficult in some settings, e.g. in primary care.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Gosden et al, 2000) **(Gosden et al, 2001)**	**Study designs** CBA, RCT, ITS **Time period** Up to 1997 (the last substantive amendment in 2000) **Participants** PCPs **Interventions** Different remuneration systems **Main outcomes** Physician behaviour	4 papers included in the Cochrane review (six in the published paper). No study assessing impacts of capitation, salary or target payments on prescribing included Only one study with an outcome relevant to prescribing (i.e. repeat prescription) (Krasnik et al, 1990). CBA, Danish study, 100 PCPs in intervention group, 326 PCPs in control The impact of adding fee-for-service to capitation on prescribing Renewal of prescription over a week period per 1000 patients significantly reduced at 12 months follow-up in the intervention group (fee for service)	No specific conclusion regarding prescribing. Concludes that FFS is likely to increase quantity of activity. The reverse was observed for repeat prescription.	"The fall in prescription renewals was a surprise since fees were introduced for them and this should encourage an increase rather than a decrease, which reduces income. This may be evidence that either PCPs do not respond to financial incentives or that the level of the fee was not sufficient to encourage such behaviour."
(Giuffrida et al, 1999)	**Study designs** CBA, RCT, ITS **Time period** 1966 – 1997 (last substantive amendment in 1999) **Participants** PCPs **Interventions** Target payment **Main outcomes** Physician behaviour, patient outcome, equity of care and costs	Only two studies were included in the review. None of the included studies meets the overview's inclusion criteria (no prescribing outcome) The use of target payments in the remuneration of PCPs was associated with improvements in immunisation rates, but the increase was statistically significant in only one of the two studies.	The evidence from the studies identified in this review is not of sufficient quality or power to obtain a clear answer to the question as to whether target payment remuneration provides a method of improving primary health care. Additional efforts should be directed in evaluating changes in physicians' remuneration systems. Although it would not be difficult to design a randomised controlled trial to evaluate the impact of such payment systems, it would be difficult politically to conduct such trials.	This systematic review's findings are included in (Gosden et al, 2001) as joint review with (Gosden et al, 2000)

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Mitchell and Sullivan, 2001)	**Study designs** Any prospective study. RCT, CCT, CBA **Setting** Primary care **Time period** 1980-1997 **Participants** PCPs and nurses **Intervention** Impacts of computers on primary care consultations **Main outcomes** Doctors performance; patient outcomes; attitudes towards computers	89 studies included. 61 studies focused on physician performance, among which 41 were controlled trials. 5 studies met the overview's inclusion criteria. In an RCT of 32 doctors [residents] and 3702 scripts computerised feedback increased generic prescribing (Gehlbach et al, 1984). This is out of the scope of the overview. A CBA study was performed to evaluate the effects of computer prompting on NSAIDs prescribing and costs (Jones et al, 1996). They observed 13% reduction in costs for the protocol group, none for computer ordering plus costs group, and increase in computer ordering only group. An RCT of 5 doctors and 658 patients concluded computerised prescribing resulted in time saving for receptionist and doctor (Roland et al, 1985). Another CBA in 25 general practices in Britain concluded that computer based decision support system did not improve lipid lowering prescribing (Hobbs et al, 1996). An RCT of 26 doctors and 5 nurses in one practice concluded that computerised reminders improved prescribing. They identified 16% increase in average physician response [prescribing and non prescribing] with reminder alone, 21% with reminder plus supporting citations (McDonald et al, 1980). The difference between these two groups was non-significant. An RCT of 681 patients in a London general practice of six GPs and one trainee concluded that computerised feedback did not change prescribing patterns for mental health patients. No difference observed between the patients' quality of life measures of the two groups after six months (Lewis et al, 1996).	"Computer use during consultations lengthened the consultation. ... Use of computers for issuing prescriptions increased prescribing of generic drugs" and reduced the costs.	It updates a previous systematic review (Sullivan and Mitchell, 1995). Unit of analysis error has not been acknowledged in this review. A study with negative findings for prescribing (Hobbs et al, 1996) has not been discussed under prescribing section of the review. The review had a subgroup for effects on prescribing. In this subgroup they only included the studies that their main outcome was prescribing. Therefore some of the studies I summarise under 'main prescribing results' in this table are not included in the original subgroup analysis. In six studies that I extracted from the review based on their design (excluding non-interventional and studies without control), four studies report improvement and two observe no change. Given the possibility of unit or analysis error in most of these studies, the review's main conclusion may be unsafe.
(Thomson O'Brien et al, 1997)	**Study designs** RCT **Setting** All settings **Time period** 1966 – March 1997. Last updated	Eighteen studies were included. Thirteen studies were on prescribing. Positive effects were observed from all studies. Nine studies had prescribing outcomes and were conducted in primary care. A Belgian study of GPs evaluated the effect of one educational visit by a specially trained GP and educational materials compared to no intervention to reduce benzodiazepine prescribing. There was 24% reduction in benzodiazepine prescribing than the control group ($p<0.05$). When outreach visits were compared to educational materials alone, the improvement was only 4% (Berings et al, 1994).	Educational outreach visits are effective in reducing inappropriate prescribing. Their cost-effectiveness is not well evaluated. When outreach visits are combined with additional interventions, they appear to be effective.	The paper briefly discusses four models of educational outreach interventions observed in the studies. The authors conclusion about the importance of preliminary interviews for the success of the outreach services has not be evaluated in the literature included in the systematic review.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	in 1997 **Participants** Health care professionals (excluding student only studies) **Intervention** Educational outreach visits **Main outcomes** Professional practice, patient outcomes (excluding knowledge, or performance in test situation)	A Swedish study attempted to improve prescribing of lipid lowering drugs for patients with hyperlipidaemia seen in primary care. For women there was 50% relative improvement in the number of prescriptions per month per health centre and a 27% relative improvement for men in the intervention group. p values could not be calculated (Diwan et al, 1995). A British study attempted to improve quality of asthma care by using clinical guidelines, educational outreach and follow up prompts. Quality of prescribing for asthma improved as a result of the intervention (Feder et al, 1995). A UK study of GPs assessed the effectiveness of a single outreach visit and educational materials compared to no intervention to encourage rational prescribing of NSAIDs. The intervention was effective (50% relative improvement, $p<0.001$), resulting in increased use of ibuprofen and reduction in average prescribing costs (Newton-Syms et al, 1992). In two trials prescribing in primary care, outreach visits were combined with social marketing: Avorn and Soumerai (1983) reduced the use of propoxyphene, cerebral and peripheral vasodilators and cephalexin in physicians identified as high prescibers. They reported a relative improvement of 15% per physician per nine months (Avorn and Soumerai, 1983). They reported that the outreach visits in addition to printed materials were cost-effective and the benefit-to-cost ratio was projected to be even greater if high prescribers were targeted (Soumerai and Avorn, 1986). A study conducted in Indonesia compared two interventions (outreach visits or a formal seminar) to no intervention to improve drug use in the management of acute diarrhoea in children. They reported that outreach visits caused a 24% relative reduction in antimicrobial use, and 40% relative reduction in the use of antidiarrhoeals compared the control. They reported that the seminar resulted in significantly greater changes than the outreach group. The use of oral rehydration agents was not significantly improved after either intervention (9% reduction). The outreach visits were less costly than the seminar ($0.77 US versus $3.30 US per participant) (Santoso et al, 1996). One trial of outreach visits plus audit and feedback reported that eight of 17 experimental providers continued to prescribe tetracycline for upper respiratory infections while 15 control providers continued to do so. The relative risk reduction was 50%, $p<0.01$ (McConnell et al, 1982).	The evidence presented here supports the use of educational outreach visits combined with additional interventions to reduce inappropriate prescribing by physicians. The cost-effectiveness of this approach in different circumstances and health care settings is unclear. The effects are small to moderate, but potentially of practical importance. It is not known how performance deteriorates over time and whether subsequent visits are cost-effective. The preliminary interviews in the social marketing approach may be important in identifying barriers to change. This process needs to be fully described so that others might replicate it. Outreach visits are costly. However, savings may outweigh costs if targeted at inappropriate prescribing and the effects are enduring (Soumerai and Avorn, 1986).	

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
		A US trial attempted to reduce the use of high cost NSAIDs (piroxicam), and increase the use of ibuprofen and salicylates by providing clinical pharmacy services in an HMO. The pharmacists provided advice to physicians and patients and reviewed patient records. The relative difference in the mean number of prescriptions of piroxicam and ibuprofen was in favour of the control group. They reported a significant difference (70% improvement) in favour of the experimental group for salicylates. The intervention was not cost-effective (Stergachis et al, 1987). An Australian RCT of GPs investigated the effectiveness of outreach visits to reduce benzodiazepine prescribing. In the primary analysis, the relative difference between the experimental and control groups was not statistically significant (Yeo et al, 1994; de Burgh et al, 1995).		
(Freemantle et al, 1997) **Study designs** RCT, CBA, ITS **Setting** All settings **Time period** Last updated in 1997. **Participants** Health care professionals **Intervention** Printed educational materials. Delivered by hand, or personal, or mass mailings. Other interventions **Main outcomes** Professional performance, patient health outcomes. Excluding knowledge, attitude, or	Eleven studies included. Four studies targeted prescribing. Only two, both RCTs, were performed in primary care. Both studies met all but two of the quality criteria set for the original systematic review (i.e. good quality). (Avorn and Soumerai, 1983) this US study reported non-significant improvement on the prescription of inappropriate medicines (3%) in the educational material group. They also observed 17.7% reduction in the prescription of targeted drugs. Given the unit of analysis error the statistical significance may be unsafe. The use of these certain products was actively discouraged by other sources and this may have augmented the effects. (Denig et al, 1990) was conducted on non-dispensing family physicians in the Netherlands and resulted in a non-significant modest reduction in the use of undesirable antispasmodics.	'Printed educational materials appear to have, at best, only a small impact on practice.' The evidence base of the two RCTs (and all others but three) may be unreliable, and this may have affected the effectiveness of intervention. Printed educational materials tend to be inexpensive. Because of lack of evidence it is difficult to establish the cost-effectiveness of printed educational materials interventions. The cost-effectiveness of additional activities that may be more effective is also uncertain.	This review has been withdrawn from the Cochrane database (checked June 2004), and seems that a new review is ongoing. The review has two comparisons. Comparison 1 is printed educational material versus no intervention. Comparison 2 is printed educational material plus another intervention versus printed educational material. The logic behind the second comparison given that the review aims to measure the effectiveness of printed educational materials is not clear. In second comparison, the difference between the intervention and control groups is in the intervention other than educational material. The comparison is beneficial in evaluating the interaction effect, and not the effects of educational material. Perhaps an alternative (or a further) comparison is between any intervention plus educational material and any intervention.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	satisfaction			Depending on the availability of evidence, this could have enabled the reviewers to assess the added effects of educational material above and beyond other interventions. Some relevant studies are not included in this review (Schaffner et al, 1983; Angunawela et al, 1991).
(Shortell et al, 1998)	**Study designs** RCT, BA, CBA, other designs **Setting** Primary care and secondary care settings **Time period** 1991 – 1997 **Participants** Health care providers **Intervention** Continuous quality improvement (CQI) **Main outcomes** Provider performance; patient outcomes	42 single site studies, of which only two randomised, were included. 13 multi-site studies, of which one randomised, one CBA, and three still in progress, were also included. Only one study met the overview's inclusion criteria. A US RCT of educational visits versus educational visits plus CQI versus none in 15 primary care clinics (97 providers) assessing prescribing for hypertension and depression among other outcomes. The educational visit plus CQI did not add any significant improvement to the intervention groups for none of the outcomes (Goldberg et al, 1998).	The observed benefits can not be confidently related to the CQI because of generally weak designs. Failures to improve the service too may be due to inappropriate choice of the target behaviour (i.e. complicated scenarios).	The authors dismissed the lack of success of the intervention in some studies, as they believed the failures were the consequences of bad implementation. The authors then explain what a perfect implementation should look like. But this may itself be a limitation of CQI: it requires vast amount of resources and energy increasing its costs, even when effective. Although the authors state that the review covers 1991-1997, at least one of the included studies in published in 1998 (Goldberg et al, 1998).

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Le Grand et al, 1999)	**Study designs** Interventional studies, with or without control. Studies of weak design were included. **Setting** Developing countries. Mainly community and primary care. **Time period** Unclear **Participants** Prescribers (physicians, health workers, pharmacists), dispensers, and consumers (patients) **Intervention** Educational, regulatory, financial, and managerial **Main outcomes** Prescribing (drugs, injections), prescribing knowledge, and beliefs	50 studies identified, only six [seven] were RCTs Four included studies with appropriate designs were out of the scope of this overview: 1 on pharmacists, two on medical students and one on health workers only. Indonesia – RCT on medical and para-medical prescribers for diarrhoea. Face-to-face education was effective in improving prescribing but 'small group face-to-face education did not have more impact than a large seminar' (Santoso et al, 1996). [Six districts, two districts randomly assigned to each group of the trial. 15 health centres included from each district. The first group received face-to-face outreach small group education plus educational material, the second group received formal seminar at the district level plus educational material, and the third group was control and received none. Education was offered only once to both intervention groups. Antibiotic use decreased by 17% in group one, 10% in group two (both significant). Anti-diarrhoeal use decreased by 8% in group one and 22% in group two (both significant). Group one improved significantly better than group two for antimicrobials, and similarly group two for anti-diarrhoeals. No change observed in control group.] Sri Lanka. Use of an antibiotic newsletter showed some improvement, but non-significant (Angunawela et al, 1991). [three-arm cluster randomised] controlled study [of 15 state health institutions (including 45 prescribers). One arm received newsletter, one arm newsletter plus educational seminar, and the third arm was control. The target behaviour was antibiotic prescribing, with the aim of reducing it. Reduction was observed in newsletter only (-7.4%), newsletter plus seminar (-7.3%), and in control (-0.4%) arm. The differences between the intervention arms and control were non significant, probably because of lack of power.] Indonesia – injections by physicians and para-medics. 70-80% of patients are seen by para-medics with no formal training (Hadiyono et al, 1996). [Before-after cluster RCT. 24 centres in one district were randomly allocated to intervention and control groups. Intervention consisted of a series of interactional group discussions of 6 prescribers and 6 patients. 27% reduction in injection in intervention, and 9% reduction in control group (significantly different in two groups). Number of drugs per prescription also reduced at the same time suggestion no substitution by other drugs.] Yemen – essential drug programme successful in improving prescribing behaviour (Hogerzeil et al, 1989; Walker et al, 1990). [Area-level controlled	Interventions such as essential drugs programmes and clinical guidelines are rarely evaluated. Most of evidence is from public sector, while irrational use of drugs is common in private sector.	Low quality studies are included. Most of the studies were relevant to non-physicians. Definition of 'prescriber' is not clear. Quality of included studies is not assessed and reported. Information included in the table within [] is obtained by examining the primary studies and not from the systematic review. Level of analysis error was unnoticed in included studies e.g. (Hogerzeil et al, 1989; Hadiyono et al, 1996; Santoso et al, 1996). Only one study involving clinical guidelines (or standard treatment guideline) was an RCT, and that one focused on pharmacists and dispensers behaviour (Ross-Degnan et al, 1996). The rest were uncontrolled before-after studies. There are at least seven RCTs included in the review and not six. Some studies that are referred to as non RCT studies in the paper, turned out to be RCTs (usually cluster randomised) after checking the original papers. Two examples are (Hadiyono et al, 1996) and (Angunawela et al, 1991). The authors admit that (Angunawela et al, 1991) is a randomised study, but seems to be suggesting it is not as good as others, because of cluster randomisation. (Hadiyono et al, 1996) is wrongly referred to as CBA study in the review. On the other hand (Hogerzeil et al, 1989)

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
		trial, data from health centres (primary care) and hospital outpatient departments included. Significant improvements were observed in the intervention group in prescribing of fewer injections (25% vs 58%), fewer antibiotics (46% vs 67%), and fewer number of total items per prescription (1.5 vs 2.4) all of which reported as significant.] [despite the reviewers' this is not a randomised trial, however it is a post-only controlled study and does not meet the overview's inclusion criteria].		is identified as a randomised controlled study, which is not true. It is a post-only controlled trial, in which a sample of health facilities in one area have been assessed only after the intervention (essential drugs programme), and compared with a sample of health centres from another area that has not received the intervention. For evaluation a random sample of health facilities in each area are chosen and the outcomes of interest are measured from them. Despite the review's authors claim, (Hogerzeil et al, 1989) is not an RCT, because the allocation of the intervention is not random. Interestingly one of the review's authors had contributed in writing the primary research! One study is reported in two sources (Hogerzeil et al, 1989; Walker et al, 1990). Although both papers are referred to in the review, only one of is identified as an included study (Hogerzeil et al, 1989), and duplication is not highlighted.
(Walton et al, 1999) (Walton et al, 2001)	**Study designs** RCT, ITS, CBA **Setting** All settings **Time period** 1966 – June 1996 **Participants**	Fifteen trials were included. The drugs studied were theophylline, warfarin, heparin, aminoglycosides, nitroprusside, lignocaine, oxytocin, fentanyl and midazolam. Interventions usually targeted doctors although some studies attempted to influence prescribing by pharmacists and nurses. All included studies took place on acute medical conditions in hospital settings. No study from primary care was identified. Computer support for drug dosage reduced the time to achieve therapeutic control, toxic drug levels, adverse reactions, and length of hospital stay.	"This systematic review provides evidence to support the use of computer assistance in determining drug dosage. Further clinical trials are necessary to determine whether the benefits seen in specialist applications can be realised in general use."	Lack of evidence from primary care may suggest lack of relevance or difficulty of use of computerised systems for this purpose in primary care.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	Health professionals **Intervention** Computerised advice on drug dosage **Main outcomes** Provider behaviour change, patient outcomes			
(Chrischilles and Gondek, 1997)	**Study designs** Different types: clinical guidelines, non-interventional studies, interventional studies **Setting** Primary or secondary care **Time period** Jan 1992 – Aug 1996 **Participants** Physicians **Intervention** Any intervention using clinical guidelines to change prescribing in a drug utilisation review context **Main outcomes** The presence and type of prescribing advice in clinical guidelines; compliance with clinical guidelines; prescribing,	56 clinical guidelines and 41 studies were included. Only 17 studies were interventions to improve care. Six [seven] studies were randomised trials. Among 17 interventional studies only five studies meet the overview's inclusion criteria (De Santis et al, 1994; Browner et al, 1994; Feder et al, 1995; Katon et al, 1995; Gorton et al, 1995). Others were excluded because of inappropriate methodology (3), setting or outcome behaviour (8), or both (1). An RCT of continuing medical education using a clinical guideline resulted in modest trend in improving the compliance in the intervention group. The result is not disintegrated in terms of prescribing and includes screening and diet management too (Browner et al, 1994). A British RCT of 39 GPs concluded that practice based education programmes using clinical guidelines can improve the quality of prescribing for the treatment of asthma (Feder et al, 1995). A CBA study concluded that a clinical guideline for treatment of asthma only improved prescribing of inhaled and oral beta-agonists (Gorton et al, 1995). Another RCT concluded that an intervention to improve dosage and duration of drug treatment for depression using a clinical guideline resulted in better care for patients in intervention group. Patient outcomes were also improved for major, and not minor, depression (Katon et al, 1995). A study using educational intervention to improve compliance with a clinical guideline recommending less use of broad spectrum antibiotics for	Clinical guidelines usually lack enough details to support DURs. Reminders at the point of prescribing are effective. (This conclusion is based on four studies, only one of them meets the overview's inclusion criteria in terms of the design and setting. This single study (Lobach and Hammond, 1994) is not on prescribing. See additional comments.) Dissemination of printed material has not been evaluated in properly designed studies. Results of group education are not impressive unless tailored to physicians' practice site (i.e. academic detailing). Studies of patient outcomes are lacking.	The search strategy, and paper retrieval was not exhaustive (e.g. the authors excluded papers because they weren't available in the local library). Also for example one included study (Lobach and Hammond, 1994) has a duplicate publication not identified by the authors (Lobach and Hammond, 1997). Although the focus of the review is on drug utilisation, the authors had included studies of preventive behaviour and adherence to immunisation guidelines. An included study (De Santis et al, 1994), which is an RCT, is not referred to as a randomised trial in the review. This paper includes primary studies were there is no specific reported primary outcome, rather papers have reported general adherence to a guideline without specifying the prescribing outcomes. Two examples within interventional studies are (Lobach

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	patient outcomes, self reported outcomes	tonsillitis resulted in some further use of penicillin V (narrow spectrum) (De Santis et al, 1994).		and Hammond, 1994) and (Browner et al, 1994). Inclusion of these studies is unsafe, as the observed improvement may have been achieved without improvement in prescribing, or despite its deterioration. More accurate checking of the full text of (Lobach and Hammond, 1994) suggests it is very unlikely that the study covered any drug prescribing at all. Other systematic reviews have not followed this systematic review's approach. For example (Mitchell and Sullivan, 2001) included (Lobach and Hammond, 1994), but did not consider it for its prescribing subgroup analysis.
(Beilby and Silagy, 1997) **Study designs** RCT, CBA, crossover, controlled time series **Setting** Primary care **Time period** 1980-1996 **Participants** General practice **Intervention** Provision of costing information **Main outcomes** Test ordering, prescribing, hospital and practice visits, and	Six studies included in the review. Two were in area of prescribing. Both were RCTs and were performed in the USA. Both studies were followed up for a period of nine months. One study meets the overview's inclusion criteria (Avorn and Soumerai, 1983). Computerised feedback on drugs increased generic prescribing, and academic detailing decreased inappropriate prescribing of target drugs. A study of 32 physicians [residents] concluded provision of monthly printouts of GPs prescribing of brand names and the potential savings if they switch to generic drugs resulted in an increase in the prescribing of generic drugs (from 14% to 67%, and to 54% at follow up) (Gehlbach et al, 1984). Increases were significantly more than control group. The participants were followed for 12 months. Academic detailing of 435 PCPs by clinical pharmacologists (average 18 min) decreased prescribing of target drugs by 13%, and resulted in a saving of $105 per physician per nine months (Soumerai and Avorn, 1986). [this paper is the economic evaluation report of (Avorn and Soumerai, 1983)]	"The provision of costing information can change GP behaviour in all service areas." Both computerised feedback on prescribing costs, and academic detailing proving these information are effective in changing prescribing. "academic detailing may be more appropriate for prescribing and computerised feedback for test ordering".	The reviewers' conclusion that computerised feedback may be more effective for test ordering and academic detailing may be more effective in prescribing is unsafe. It is not supported by evidence provided in the review. The study could have included (Harris et al, 1984), which had been included in a previous systematic review (Mugford et al, 1991).

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	costs			
(Anderson and Lexchin, 1996)	**Study designs** RCT **Setting** Primary care **Time period** 1989 - 1995 **Participants** Physicians **Intervention** Different interventions targeting prescribing **Main outcomes** Prescribing	Nine studies were included. Two studies were on interns or residents, three in outpatient clinics and one in hospital. A US trial, conducted in an HMO used a 10-minute, face-to-face educational meeting between prescribers and a pharmacist to achieve a more appropriate use of anti-ulcer medications. Differences between the control and intervention groups were significant after one month, but disappeared by the second month after the intervention (Raisch et al, 1990). A British study targeted prescribing of NSAIDs in an intervention involving visits by a specially trained pharmacist. Each community GP in the intervention group received a single visit. The intervention group showed an improvement in the appropriateness of prescribing that lasted for at least five months (Newton-Syms et al, 1992). An Australian study showed that a group of community GPs who received mailed educational materials, followed by a visit from a pharmacist, had better compliance rates than a control group for prescribing antibiotics in tonsillitis (De Santis et al, 1994). Another US study compared a group of physicians who received feedback (involving chart review, two letters and individualized suggestions for reducing poly-pharmacy), a group that received a single letter identifying patients who had been given ten or more prescriptions and recommending a reduction in the number of medications, and a control group that received no intervention. Poly-pharmacy was reduced in both intervention groups compared with the control group, but there was no statistically significant difference between the effects of the two interventions (Meyer et al, 1991) [in hospital outpatient].	The existing research indicates that the dissemination of printed material alone does not lead to improvements in practice, but specific educational and feedback strategies can improve the quality of care. Aside from trials using academic detailing, nearly all published studies have been conducted in academic family practice units, large group practices or outpatient clinics. Only one study (in hospital setting) looked at a broader range of quality-of-care issues, including inappropriate choice of therapy, under-dosage and over-dosage, scheduling, drug-drug interactions, therapeutic duplication, allergies and the omission of necessary drug therapies.	It includes non-interventional studies: current sources of information for Canadian community-based physicians. The study covers the period after the publication of (Soumerai et al, 1989). The authors report the findings of other systematic reviews under the heading 'findings from the literature review', along side what they extracted from the primary research. This can potentially be misleading, given that they do not provide any evaluation of the message from other reviews and their validity.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Gill et al, 1999)	**Study designs** RCT, CBA, CCT **Setting** Any **Time period** 1975 - 1994 **Participants** Doctors (including in-training) **Intervention** Different professional interventions **Main outcomes** Those reported in primary studies. Categorised as positive, negative, or inconclusive	79 studies including 96 interventions. 25 studies (33 interventions) were conducted in the community (primary care). The counting was repeated among the 33 interventions to improve prescribing in primary care for this overview based on the data provided in the tables: Studies using audit and feedback (5/5), and patient mediated interventions (4/4) showed significant improvement in 100% of circumstances. Multi-faceted interventions were successful in 53% (10/19), educational materials in 50% (2/4), and outreach in 0% (0/1) of circumstances.	No separate analysis is reported for primary care studies. Here are the general findings of the review: 49 (51%) showed significant change in comparison to the intervention. Interpretation for specific interventions is not possible because of wide variation. Patient mediated interventions are the most successful (63%) followed by audit and feedback (52%), and outreach (50%). Multi-faceted interventions were successful in only 49% of comparisons. Educational materials where used alone is less likely than other interventions to be effective.	In their final categorisation of studies the authors have pooled together some categories that others may consider separately: educational outreach with opinion leaders; and audit and feedback with reminders. The systematic review summarises the evidence using a simple approach (vote counting): giving '+' to each study with statistically significant positive findings, '=' to those with non-significant difference, and '–' to those with statistically significant negative results. They use z transformation for quantitative meta-analysis. The authors include vaccination prescribing in the review. When I repeated the vote counting in primary care studies only, educational materials performed as well as multi-faceted interventions and better than outreach.
(Balas et al, 1996)	**Study designs** RCT **Setting** All settings **Time period** Not stated **Participants** Physicians	12 studies were included in the review. Only one study meets the overview's inclusion criteria. A Danish RCT concluded that peer-comparison feedback reduced prescribing levels pre practice (Lassen and Kristense, 1992).	Peer-comparison feedback (profiling) has a minimal but significant effect on utilisation. The subgroup analysis for prescribing yields similar results but lack statistical power. The two prescribing studies included in this analysis were conducted in outpatient clinics (not primary care). The cost-saving effects of feedback are unlikely to exceed costs of the intervention. (no evidence is provided for this, but it is based on	The study uses vote counting and z transformation and odds-ratio techniques for quantitative meta-analysis. The conclusions are unsafe given significant heterogeneity observed between the included studies. One included study is not an RCT, rather a CBA study (Lassen and Kristense, 1992).

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
			small effects observed)	
	Intervention Peer comparison feedback (physician profiling) **Main outcomes** Utilisation of clinical procedures			
(Grimshaw et al, 1995) **(NHS Centre for Reviews and Disseminatio n, 1994)**	**Study designs** BIB, RCT, CBA, ITS **Setting** All settings **Time period** 1975 – June 1994 **Participants** Physicians **Intervention** Interventions to implement clinical guidelines **Main outcomes** Provider performance; patient outcomes	91 studies were included, of which 22 are studies of prescribing, the use of radiological or laboratory investigations. Eight studies were categorised as prescribing studies, among which three were in primary care. At least one of the studies in clinical care group was relevant to the overview's objectives. A US time series study evaluated the implementation of state guidelines to reduce injectable antibiotic prescribing for respiratory illnesses. Distribution by targeted mailing, practice visits, and financial incentives (payment denial unless complied with the guidelines) were used for implementation. The intervention was effective and antibiotic injections fell by 60% while other prescriptions hardly changed (Brook and Williams, 1976; Lohr and Brook, 1980). A US CBA studies successfully used educational outreach to implement clinical guidelines to reduce long term use of diazepam (Ray et al, 1986). Another US CBA study used educational outreach [and based on the review, unsuccessfully] to implement clinical guidelines for improving anti-ulcer treatment (Raisch et al, 1990). Checking the original paper, the intervention groups performed statistically better in one month after the intervention, and they were still better, but not anymore statistically significant, two month after the intervention. A British BIB study of guideline implementation for five common paediatric conditions concluded that participatory guideline development improved prescribing for all five conditions by 8% on average. The groups receiving the guidelines did not improve as such (North of England Study of Standards and Performance in General Practice, 1992a; North of England Study of Standards and Performance in General Practice, 1992b).	Of 87 studies examining the process of care 81 reported significant improvements. Clinical guidelines can improve clinical practice and may lead to improved patient outcomes. There is some evidence of the effectiveness of the educational outreach in improving prescribing. Educational interventions that require more efforts are more costly and their effectiveness should be assessed. Mailed educational materials (clinical guidelines) are occasionally effective in changing behaviour, e.g. radiological investigations). Educational activities related to clinical decision making are more likely to be effective, but not enough evidence to know relative effectiveness of different interventions in different contexts.	The reviews update a previous systematic review (Grimshaw and Russell, 1993). The systematic review distinguishes between clinical care studies which may or may not have any prescribing outcome measure, and prescribing studies in which prescribing outcomes are the main outcomes. For the former groups of studies (clinical care), the collective effect on clinical care is reported in the systematic review. Study selection, quality assessment and data extraction processes are not explained.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Wensing and Grol, 1994) **and its update:** **(Wensing et al, 1998)**	**Study designs** RCT, CBA, other designs **Setting** Primary care **Time period** 1980 – 1992 (1980 – June 1994 for the update) **Participants** PCPs **Intervention** Different interventions for guideline implementation, with emphasis on the difference between single and combined interventions **Main outcomes** Professional performance	75 studies were included in the original review. 27 studies were RCTs and 17 were of CBA designs. The update review includes 143 studies, of which 39 were RCTs and 22 were CBAs. Single strategies: A CBA study concluded that tutorial sessions were effective in promotion of appropriate antibiotic prescribing for UTI (Klein et al, 1981). Two CBA of educational outreach were effective in reducing long term diazepam prescribing (Ray et al, 1986) and contradicted antibiotics prescribing (Schaffner et al, 1983). Multi-strategy: Educational outreach + material: an RCT found them effective in improving prescribing behaviour (Soumerai and Avorn, 1986). A CBA of feedback + peer review observed less increase in prescribing charges in the intervention group (Harris et al, 1985). Another CBA observed non-significant decrease in prescribing cost per patient per DDD owing to feedback + peer review (Lassen and Kristense, 1992) In a British CBA peer review (standard setting) + feedback in two of five standard setting groups children were prescribed less antibiotics in the fist visit, and in one group less of other drugs (North of England Study of Standards and Performance in General Practice, 1992a). Another CBA in Mexico focused on improving prescribing for acute diarrhoea (Gutiérrez et al, 1994). RCT of implementing dyspepsia clinical guidelines (Jones et al, 1993)	Individual instruction, feedback, and reminders are the most successful interventions. The most effective combined interventions are all combinations of individual instruction, and the combination of peer review and feedback. It suggests that combined feedback and reminder is not an effective combination. Combination of performance oriented strategies may not be that effective. Conclusions from (Wensing et al, 1998): Reminders are often used for influencing prevention and peer review groups for changing therapeutics. Single interventions using feedback were more often effective than combinations of feedback and information transfer. Information transfer is probably always needed at some point in the process of implementing change. Some, but not all, multi-faceted interventions are effective in inducing change in general practice.	The original systematic review (Wensing and Grol, 1994) has been updated by the same group of researchers (Wensing et al, 1998). Unfortunately the updated study does not link individual conclusions and findings to the primary studies that provide the evidence for that conclusion. Hence both reviews are discussed together in here. Original review's search strategy was not exhaustive, especially for prescribing. The search for the update review is improved. None of the RCTs involving a single intervention were eligible for consideration in the overview. The studies that targeted prescribing were conducted in community hospital, or outpatient clinics, or involved residents as the participants. The approach use in the systematic review to consider which combinations for multi-faceted interventions are more likely to be effective is promising. Unlike many other reviews that just suggest multi-faceted interventions are more effective, here they try to identify effective combinations. This is also supported by their conceptual mapping of interventions. Definitions used, especially in the original review, are not very clear. The review uses 'face-to-face instruction' or 'individual

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
				instruction' instead of educational outreach. Also it is not clear what exactly the reviewers mean by 'peer review' or 'practical support'. Some definitions for the terms are provided in another paper by the same group of researchers (Grol, 1992). The North of England study (1992) is a BIB and not a CBA.
(Yano et al, 1995)	**Study designs** Different designs, if meet certain criteria **Setting** Primary care **Time period** January 1980 – December 1992 **Participants** Physicians in primary care setting and outpatient clinics **Intervention** Any intervention **Main outcomes** Provider performance, patient outcome, utilisation costs	36 articles were included. Three studies meet the overview's inclusion criteria. In two separate studies, educational outreach visits by physician counsellors reduced long term use of diazepam, and the use of contradicted antibiotics and expensive oral cephalosporins (Ray et al, 1985b; Ray et al, 1986). Notification of doctors of elderly patients with 10 or more number of medications (mail reminders) reduced the average number of medications per patient from 12 to 6 (Meyer et al, 1991). [outpatient clinic] Educational outreach visits by pharmacists reduced inappropriate drug prescribing (Soumerai and Avorn, 1986; Soumerai and Avorn, 1987).	A set of recommendation on how primary care can achieve its objectives, as defined in the paper.	Search strategy does not seem to be exhaustive. Study starts with a short review of primary care objectives. Study inclusion criteria both in terms of outcomes and interventions are wide. No recommendation on prescribing is being provided. The main paper of one of included studies (pharmacist-led educational outreach) was not referred to in the systematic review (Avorn and Soumerai, 1983).

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Pippalla et al, 1995)	**Study designs** RCT, CBA **Setting** Primary, secondary, or community **Time period** Jan 1979 – Sep 1991 **Participants** Physicians **Intervention** Interventions to influence prescribing **Main outcomes** Professional behaviour	26 studies were included. After checking the full-text or abstracts of the papers, nineteen studies did not meet the overview's inclusion criteria: laboratory test ordering (3), hospital setting – inpatient or outpatient (10), residents or medical students (2), nursing home (1), repeated publication of one study (2), and irrelevant study (1). Seven studies meet the overview's inclusion criteria: (Wang et al, 1979; Klein et al, 1981; McConnell et al, 1982; Avorn and Soumerai, 1983; Schaffner et al, 1983; Stergachis et al, 1987; Denig et al, 1990). For only one of these studies any information other than citation and the effect size (as a number) is provided in the paper: Printed educational material of cost information and one-to-one meetings had a noticeable and significant effect on antibiotic prescribing (Klein et al, 1981).	Interventions targeting prescribing behaviour are effective in changing prescribing (based on statistical summary. Mean Z=0.56). Mean Z (effect size) for drug prescribing studies was 0.54. Community (primary care) physicians are more prone to change in their prescribing than others (mean Z=0.71). Prescribing restrictions by formularies are the most effective interventions (mean Z=0.93). In terms of interventions 'one-to-one' meetings were the most effective (0.64) and group lectures were the least effective (0.31) interventions.	This paper updated (Raisch, 1990a) and (Raisch, 1990b). The update was done without systematic literature search, hence prone to selection bias. The pooled results should be used very cautiously since the effect sizes from included studies were significantly heterogeneous ($p \leq 0.0001$). Laboratory test ordering and overall medical care were also included in this review. For most readers prescribing behaviour does not include laboratory test ordering. Despite this, there was no mention of inclusion of this group of studies in the title or the abstract of the systematic review. The paper did not report which papers were included for each sub-group analysis. Therefore it was uncertain whether the effect sizes reported for each of these sub-groups were relevant to drug prescribing only, or included other behaviours. Also it was not possible to separate results from primary care and other settings. Three of the included papers reported the results of one RCT from different angles (re-analyses) (Avorn and Soumerai, 1983; Soumerai and Avorn, 1986; Soumerai and Avorn, 1987). One included study was not an intervention to change behaviour, and had no behaviour outcome.

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
				Rather the investigators had provided clinicians and patients with different levels of information about lung cancer and measured their preferences between radical surgery and radiotherapy (McNeil et al, 1982). The paper should not have been included.
(Mugford et al, 1991)	**Study designs** BA, CBA, RCT **Setting** Primary care, secondary care, other settings **Time period** Not stated **Participants** Health care professionals **Intervention** Effects of feedback of statistical information on behaviour change **Main outcomes** Clinical practice	36 studies were included, of which eight were CBA studies and twelve were RCTs. Seven studies were focused on prescribing. Five were controlled studies. One study was conducted in hospital setting. Another controlled study had teaching practices in the intervention group and non-teaching practices in the control, and hence considered as not meeting the overview's inclusion criteria. A further study was on monitoring of patients receiving a drug rather than prescribing itself. Two studies meet the overview's inclusion criteria. One US CBA of three interventions found no effect from mailed brochures on antibiotic prescribing (Schaffner et al, 1983). A British CBA study provided feedback (twice a year for two years) to GPs in the intervention group about number and cost of prescription as well as meetings for discussion. The intervention was carried out at the practice level, as well as on individual GPs. The intervention group prescription rate per patient fell more than the control group (Harris et al, 1984).	Information feedback is most likely to change behaviour if decision makers have already agreed to take part. It is more likely to be effective if the information is provided close to the time of decision making.	(Jamtvedt et al, 2003) questioned the validity of the review's conclusion that audit and feedback close to the time of decision making can be more effective. It is not clear why (Schaffner et al, 1983) was included in this systematic review. There is no audit and feedback intervention in this trial, and the mail brochure arm did not incorporate any audit or feedback data. None of the five primary studies included in this systematic review that were on prescribing behaviour and were controlled clinical trials were included in (or even considered and then excluded from) a later EPOC Cochrane Review Group systematic review of the effectiveness of audit and feedback (Jamtvedt et al, 2003). Given that two of these primary studies had 'audit' or 'feedback' in their title, this is surprising (Anderson et al, 1988; Parrino, 1989).
(Soumerai et al, 1989)	**Study designs** ITS,CBA, RCT, as well as uncontrolled, and post only designs	44 studies were included. Dissemination of printed educational material: Mail print material: two studies. One large RCT in the US identified a small	Analysis of the best controlled studies failed to support that mailed printed material when used alone could effectively change physicians' practice. It is possible though that	The definition of ambulatory setting used in this systematic review is not equivalent with what is used in the overview of systematic reviews. This systematic review

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
	Setting Ambulatory care (includes emergency rooms, and hospital outpatients) **Time period** 1970-1988 **Participants** Physicians and in-training physicians **Intervention** Different interventions targeting prescribing in ambulatory care **Main outcomes** Clinical behaviour outcomes, patient outcomes	non-significant reduction (-4%) in inappropriate prescribing in the mail only intervention group comparing to the control (Avorn and Soumerai, 1983), which may be cost effective (Soumerai and Avorn, 1986). Another US CBA study (four arms) oncluded that an attractively designed mail brochures did not reduce inappropriate prescribing of any of four different antibiotics in comparison to the control group (Schaffner et al, 1983; Ray et al, 1985a). Protocols and guidelines: In a US CBA study, participatory developed guideline of prescription and over-the-counter drugs improved prescribing in the development site, but not in another site that received the guideline in comparison to the control. The effect in the development site was more due to the adherence of those physicians who were directly involved in the guideline development (Bush et al, 1979). Self-instruction materials. A Canadian RCT provided 14 weekly mailing packets of self-study material on consequences of improper management of hypertension. No change was observed in physicians' long-term knowledge, practice, or patients' outcomes (Evans et al, 1986). A factorial design RCT of two interventions (4 arms) assessed the impacts of computerised audit and self-education material on hypertension management. The interventions resulted in more frequent visits by patients to the doctors offices, but patients' outcomes did not improve for none of the intervention groups comparing to the control (Dickinson et al, 1981). Mailed national warning campaigns. A six-year spanned Swedish ITS concluded that letter warnings of dangers of using chloramphenicol and dipyrone as well as publications in the Journal of Swedish Medical Association may have resulted in 60% reduction in the use of the drugs, and 80% reduction in their adverse consequences (Bottiger and Westerholm, 1973). Two British ITS studies also recorded reduction in the use of pressurised aerosols (for asthmatics) and chloramphenicol (Inman and Adelstein, 1969; Wade and Hooh, 1972). A US ITS study found FDA warnings, labelling changes and press releases ineffective in changing propoxyphene prescribing patterns (Soumerai et al, 1987a). A Northern Ireland study also demonstrated a downward trend in barbiturate use but it was probably inline with the secular trend observed before the campaign (King et al, 1980). Reporting of patient specific listings of prescribed drugs: Three studies were conducted in outpatient clinics. An RCT reported above evaluated patient profiles with or without educational material. No	mailed materials are cost-effective owing to low costs. Well designed educational materials are an important component of other interventions such as educational outreach and feedback. Nationwide warning programmes without media campaigning may not be effective. Merely reporting detailed drug use profiles of individual patients is unlikely to improve prescribing of busy [primary care] physicians. Group educational interventions may change attitudes and knowledge, but have unclear effects on practice. The conclusion is based on academic settings. Ongoing feedback from credible sources can increase generic prescribing and adherence to protocols. No well-controlled studies had been reported from less organised office-practice setting. Ongoing computerised reminder systems could be effective in preventing physicians from omitting essential measures for several diseases. They correct errors of omission, and act as secretarial reminders. Face-to-face educational outreach is effective and cost-effective in changing prescribing behaviour. Inadequately controlled studies are	includes emergency rooms, and even community hospitals in their analyses, e.g. (Stross and Bole, 1980). (McConnell et al, 1982) has been included in other studies and discussed in details. One systematic review did not mention the huge baseline differences between the intervention and control groups that could invalidate the findings of the study (Jamtvedt et al, 2003). A duplicate publication of this review exists (Soumerai et al, 1990).

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
		effect was observed in comparison with the control (Dickinson et al, 1981). Group education. Three study were included. Two were in outpatient clinics and one in inpatient wards. Feedback of prescribing patterns. Two controlled studies included: one on family practice residents (Gehlbach et al, 1984) and the other on house officers. Reminders at the time of prescribing. Two studies on monitoring or pre-treatment work ups, one on interns and residents, one in outpatient, and one on community surgeons. Four years of time series data demonstrated marked reduction in the percentage of untreated streptococcal pharyngitis by antibiotic as a result of computer generated reminders for doctors or nurses. The effect disappeared as soon as the reminders stopped (Barnett et al, 1978). Face-to-face educational outreach. An RCT evaluated provision of two educational outreach plus educational printed material by clinical pharmacists without prescribing feedback improved (reduced) prescribing of three drugs (Avorn and Soumerai, 1983). The effects were regardless of physicians' background characteristics. The second follow up outreach visit was highly correlated with prescribing change (Soumerai and Avorn, 1987). The intervention was highly cost-effective (benefit/cost ratio of 2.0 or higher) (Soumerai and Avorn, 1986). Another US CBA study concluded that 15 minutes outreach visits (plus printed material) by physician counsellors was effective in reducing inappropriate antibiotic prescribing, while outreach visits by a pharmacist drug-educator was not (Schaffner et al, 1983). The improvements as a result of the intervention were persistent in the second year of follow up although the effect was smaller (Ray et al, 1985b). Another analysis of this experiment found no effect from physician counsellors outreach visits on diazepam prescribing (Ray et al, 1986). Another US RCT used physician counsellor outreach visits (plus individualised feedbacks) to reduce tetracycline prescribing for viral respiratory infections. The baseline rate of prescribing in the intervention group was twice as the control group, questioning the randomisation process. It was effective in reducing tetracycline prescribing, but could be subject to biases (McConnell et al, 1982). Clinical pharmacy services. One in outpatient, and the other study of very poor design.	more likely to report positive findings than adequately controlled studies.	

Authors	Characteristics	Main prescribing results	Review's main conclusions	Additional comments
(Lomas and Haynes, 1988)	**Study designs** RCT **Setting** All settings **Time period** 1975-1987 **Participants** Physicians (in practice or training) **Intervention** Different interventions **Main outcomes** Physician performance, patient outcome	32 studies were included. Six studies meet the review's inclusion criteria. Educational strategies Mailed educational material. No effect from educational material on prescribing for three different drugs (Avorn and Soumerai, 1983). Mailed weekly educational materials on the management of hypertension had no effect (Evans et al, 1986). Individually tailored instruction. Three studies included. No specific prescribing outcome reported. Educationally influentials and preceptorships. Two studies both conducted in community hospitals. Traditional quality assurance measures. Seven studies were included: one in hospital, one on neonatal care, one in emergency room, two on patient care appraisal or quality assurance for several items, and one on the use pelvimetry. One study was focused on prescribing and concluded that outreach services are effective in improving prescribing (Avorn and Soumerai, 1983). Expert computer programmes. Three studies were included. All three were conducted in hospital settings. Computerised medical records. Four studies were included, none on prescribing in primary care. Computer-aided audit and feedback for cost containment. Three studies were included. All three were conducted on residents. Computer-aided audit and feedback mechanism for quality assurance. Eight studies were included. Four studies were relevant to prescribing in ambulatory care and showed some effect in changing physician behaviour.	"the resources currently being spent on the production and distribution of such materials [mailed educational materials] can be diverted into more effective approaches" p87. Strategies such as mailing materials and didactics course if are used at all, they be incorporated into other programmes. Strategies targeting multiple determinants are more likely to be successful. Innovative strategies should incorporate elements of different categories of interventions.	The last part of the paper is a systematic review of interventions, updating two previous reviews (Haynes et al, 1984; Haynes and Walker, 1987). The review is more focused on preventive care behaviours, but covers all types of behaviours. Categorisation of the interventions is not that useful. For example inclusion of a trial of educational outreach (Avorn and Soumerai, 1983) under the title of traditional quality assurance is not that helpful. On the other hand several categories are given to computer related interventions. The systematic review is too dismissive regarding mailed educational materials, and it that sense it goes beyond the message from one of the high quality included studies (Avorn and Soumerai, 1983). In their economic evaluation, (Soumerai and Avorn, 1986) suggested that mailed educational materials can be cost-effective. It is also much more positive in its tone to the use of computerised systems than the systematic review of prescribing (Soumerai et al, 1989). This may be due to the reviews' different points of focus (all behaviours with emphasis on preventive care versus prescribing), but is more likely to be due to the way the evidence is being interpreted by the authors.

* Information provided between brackest [ ] is mostly extracted from primary studies (and not the reviews).

Appendix II-3. Discarded papers from overview of systematic reviews of interventions to improve prescribing in primary care

(ascending in publication order*)

Reference	Focus of the review	Reason for discarding / further information
(Haynes et al, 1984)	Continuing medical education	The review has been updated in (Lomas and Haynes, 1988; Davis et al, 1992; Davis et al, 1995; Davis et al, 1999)
(Soumerai and Avorn, 1984)	Improving prescribing in hospital setting	Focusing on prescribing in secondary care.
(Horder et al, 1986)	Changing GPs' behaviour	Not a systematic review.
(Haynes and Walker, 1987)	Quality improvement using computers	The review has been updated in (Lomas and Haynes, 1988; Johnston et al, 1994)
(Schroeder, 1987)	Interventions to reduce medical costs by changing physicians' behaviour	It is a non-systematic review of evidence.
(Beaudry, 1989)	Continuing medical education	No focus on prescribing. The review does not provide the reference list of the included studies.
(Gurwitz et al, 1990)	Improving prescribing in nursing homes	Focusing on prescribing in nursing homes.
(Raisch, 1990a)	Changing physicians' prescribing	It has been updated in (Pippalla et al, 1995)
(Raisch, 1990b)	Changing physicians' prescribing	It has been updated in (Pippalla et al, 1995)
(Lomas, 1991)	Effects of consensus recommendations	Prescribing not discussed. It concludes Dutch studies show more success than North American studies.
(Waddell, 1991)	Continuing education	Focus on nursing practice
(Davis et al, 1992)	Continuing medical education	The review includes several interventions other than CME. It has been updated in many later reviews, including (Davis et al, 1995).
(Grol, 1992)	Implementing guidelines in general practice	This is not a systematic review of primary literature
(Long and Sheldon, 1992)	Improving purchaser and provider decisions	Not a systematic review.

Reference	Focus of the review	Reason for discarding / further information
(Axt-Adam et al, 1993)	Changing physicians' behaviour	No prescribing. Lab test ordering
(Buntinx et al, 1993)	Feedback and reminder in ambulatory care	No prescribing. Preventive and diagnostic behaviour
(Greco and Eisenberg, 1993)	Changing physicians' behaviour	Non-systematic review of evidence. Not methods are presented.
(Grimshaw and Russell, 1993)	Clinical guideline implementation	It has been updated in (NHS Centre for Reviews and Dissemination, 1994) and (Grimshaw et al, 1995)
(Kreling and Mott, 1993)	Cost effectiveness of DUR programmes	No study met the reviewers' inclusion criteria
(Ackermann and Cheal, 1994)	Physician's adherence to guidelines	No prescribing. Breast cancer screening. Surveys were included.
(Austin et al, 1994)	Effectiveness of physicians' reminders	No prescribing. Preventive care
(Gyorkos et al, 1994)	Effectiveness of immunisation delivery methods	No prescribing. Immunisation programmes
(Haines and Jones, 1994)	Evidence implementation	Not a systematic review
(Johnston et al, 1994)	Improving quality of care by computer-based CDSS	It has been updated in (Hunt et al, 1998)
(Silagy et al, 1994)	Effectiveness of health professionals training	No prescribing. Smoking cessation interventions.
(Conroy and Shannon, 1995)	Clinical guideline implementation in primary care	Not a systematic review. More like an overview of SRs. Very little on prescribing.
(Davis et al, 1995)	Continuing medical education	The review covers interventions other than formal CME. It has been updated in several later systematic reviews.
(Mandelblatt and Kanetsky, 1995)	Changing physicians' behaviour	No prescribing. Breast cancer screening
(Oxman et al, 1995)	Interventions to improve professional practice	The systematic review has been updated in a few systematic review (including EPOC systematic reviews) published after it.
(Sullivan and Mitchell, 1995)	Effects of computers on primary care consultation	The review has been updated in (Mitchell and Sullivan, 2001)
(Bloor and Freemantle, 1996)	Controlling pharmaceutical expenditure through changing physician behaviour	Not a systematic review
(Bloor et al, 1996)	Controlling pharmaceutical expenditure (regulating industry)	Not a systematic review
(Freemantle, 1996)	Effective interventions for	Not a systematic review

Reference	Focus of the review	Reason for discarding / further information
	decision making	
(Freemantle and Bloor, 1996)	Controlling pharmaceutical expenditure through targeting patients	Not a systematic review
(Kerwick and Jones, 1996)	Changing physicians' behaviour in primary care	A non-systematic review. Focus on psychiatry.
(Shea et al, 1996)	Improving preventive care in ambulatory setting	No focus on prescribing.
(Snell and Buck, 1996)	Improving cancer screening	No focus on prescribing.
(Davis and Taylor-Vaisey, 1997)	Clinical guideline implementation	Not focused on specific physicians' behaviours (e.g. prescribing) or settings (e.g. primary care)
(Hanson et al, 1997)	Improving care at the end of life	Very specific target group
(Howard and Duncan, 1997)	Evidence implementation	Not a systematic review. Prevention of atrial fibrillation
(Hulscher et al, 1997)	Improving preventive care in primary care	No focus on prescribing. Last updated in 1996. Possibly linked to (Hulscher, 1998)
(Worrall et al, 1997)	Clinical guideline implementation in primary care	No focus on prescribing
(Bero et al, 1998)	Evidence implementation	Not a systematic review of primary research. It is an overview of systematic reviews. It has been updated in (NHS Centre for Reviews and Dissemination, 1999)
(Dowie, 1998)	Clinical guideline implementation in the UK	Not a systematic review of published evidence. It focuses on what research initiatives are happening in the UK.
(Freemantle et al, 1998)	Methodology of evidence implementation studies	No prescribing
(Garner et al, 1998)	Evidence implementation in developing countries	Not a systematic review
(Grimshaw, 1998)	Improving out-patient referrals	No focus on prescribing
(Hulscher, 1998)	Preventive care in general practice	This study was not assessed because of access. Possibly linked to (Hulscher et al, 1997)
(Smith et al, 1998a)	Continuing medical education in primary care	Not a systematic review of primary studies. It reports the results of two systematic reviews already included in the overview.
(Solomon et al, 1998)	Improving use of diagnostic procedures	No focus on prescribing

Reference	Focus of the review	Reason for discarding / further information
(Thomas et al, 1998)	Clinical guideline implementation	Not on physician behaviour. Focus on professions allied to medicine
(Cabana et al, 1999)	Clinical guideline implementation	No focus on prescribing. Broad inclusion criteria
(Cantillon and Jones, 1999)	Continuing medical education in general practice	It is an overview of systematic reviews and other studies.
(Davis et al, 1999)	Continuing medical education	No discussion of prescribing. An update of (Davis et al, 1995). It has been updated in (Thomson O'Brien et al, 2001)
(Durieux et al, 1999)	Continuing medical education	Not addressing general practice. In French
(Freudenstein and Howe, 1999)	Continuing medical education in primary care	No focus on prescribing. No discussion or conclusion about it in the paper. Only in table of included studies, prescribing has been mentioned. There is no data synthesis for prescribing
(Gosden et al, 1999)	Impacts of salary payments on physician behaviour	The review has been updated in (Gosden et al, 2000)
(Greenhalgh and Meadows, 1999)	Changing physicians' behaviour	No focus on prescribing. Review of use of patient based measures of health
(NHS Centre for Reviews and Dissemination, 1999)	Evidence implementation	Not a systematic review of primary research. It is an overview of systematic reviews. It is an update of (Bero et al, 1998)
(Pagliari and Kahan, 1999)	Clinical guideline implementation	No focus on prescribing. Barriers and attitudes.
(Shiffman et al, 1999)	Computer-based guideline implementation	No data synthesis in prescribing or primary care. Very superficial coverage of prescribing in the table of included studies.
(Thomas et al, 1999)	Clinical guideline implementation	Not on physician behaviour. Focus on professions allied to medicine
(Balas et al, 2000)	Physicians' prompting	No focus on prescribing. Focuses on preventive care.
(Durieux et al, 2000)	Clinical guideline implementation	It is an overview of systematic reviews. The first author wrote to me that it does not address GPs' behaviour specifically. In French
(Grol and Jones, 2000)	Evidence implementation	Not a systematic review.
(Johnston et al, 2000)	Barriers and facilitators for audit	No focus on prescribing. Studies of physicians' and non-physicians' providers are included. Research methods other than those accepted by the EPOC are included.

Reference	Focus of the review	Reason for discarding / further Information
(Lancaster et al, 2000)	Improving smoking cessation in primary care	Focused on smoking cessation advice and brief consultations. Last updated in 2000.
(Ockene and Zapka, 2000)	Provider education for clinical guideline implementation	Not a systematic review. The methods are not presented.
(Smith, 2000)	Changing physicians' behaviour	Not a systematic review. It is an overview of systematic reviews.
(Thomson O'Brien et al, 2000)	Audit and feedback versus alternative interventions	Withdrawn from the Cochrane Library as it is now covered by (Jamtvedt et al, 2003). Only one prescribing paper included in the review, which was on interns in secondary care setting.
(Grimshaw et al, 2001)	Changing physicians' behaviour	It is an overview of systematic reviews.
(Stone et al, 2002)	Improving adult immunisation and cancer screening in ambulatory care	No focus on prescribing.

Appendix IV. Semi-structured interviews' plan

a) Summarising statement about the project and the purpose of the interviews by the interviewer.

b) Clarifying that the interview will be tape-recorded (consent to record)

c) Questions: (use questions as a guide only)

1) Thinking of clinical guidelines in primary care, can you tell me of examples where the guidelines were useful in drug prescription of general practitioners?

   a. *Probe: asking for specific guidelines*
   b. *Why do you think the guidelines were useful?*

2) Thinking of clinical guidelines in primary care, can you tell me of examples where the guidelines were not useful in drug prescription of general practitioners?

   a. *Probe: asking for specific guidelines*
   b. *Why do you think the guidelines were not useful?*

The next set of questions asks about your opinions on clinical guidelines that have been developed for specific clinical conditions.

3) **Asthma:** Among clinical guidelines for the treatment of asthma in adult patients, which guidelines are more likely to be followed in general practice?

   a. Why do you think some guidelines were adopted in general practice rather than others? How those guidelines influenced prescribing for asthma?

4) **Statins:** How do you think of the advantages of clinical guidelines for the use of statins? What are the possible results of adherence to guidelines for the use of statins?

5) **Depression:** Are you familiar with any guideline for the treatment of depression in primary care? Is it likely to be followed in general practice? Why?

6) **Epilepsy:** Are there any clinical guidelines that you use in the treatment of epilepsy? Why?

7) **Menorrhagia:** Which guidelines, if any, do you use in the treatment of menorrhagia? What influences has it made on your prescribing patterns?

8) Who else might influence GPs' adherence to any of the guidelines, which have been mentioned in the interview so far?

    a. *Probe: does it differ for different guidelines?*

9) **Barriers:** What are the barriers of the clinical guidelines' implementation in your practice?

    a. *The guidelines can mentioned specifically again, if necessary*

10) **Facilitators:** What do you think are the factors that encourage you to use those clinical guidelines in every-day practice?

11) Is there anything else I should have asked you?

THE UNIVERSITY of York

Department of Health Sciences

National Survey of General Practitioners
Perceptions and Practice of Statin Prescribing

WE SHALL KEEP ALL RESPONSES **STRICTLY CONFIDENTIAL**

We want to analyse your responses in the context of your prescribing practice. We should therefore be most grateful for permission to use your Prescribing And CosT [PACT) data on asthma and statin prescribing for one year from last quarter. Please sign here if you are willing to give access to those data solely for this research.

1. Do you permit us to access PACT data on your personal prescribing for asthma and of statins? Yes ☐ No ☐
2. Are you the senior partner in the practice? Yes ☐ No ☐
3. If yes, do you permit us to access PACT data for your practice? Yes ☐ No ☐

Signature ..... Date .....

Please fill in the questionnaire even if you do not wish us to access your PACT data.

****

We shall include all returned questionnaires in a lottery.
The winner will receive a **digital camera** early next year.
This is a token of our gratitude for your valued contribution.

I. ABOUT YOU AND YOUR PRACTICE *Please tick whichever applies or write the answer where applicable*

1. Age: ..... 2. Sex: F ☐ M ☐ 3. Year of graduation from medical school? .....

4. Are you:
   Principal in practice Yes ☐ No ☐
   Salaried GP (PMS contract) Yes ☐ No ☐
   Other (specify) .....

5. Practice status: *(please tick all that apply)*
   a. Dispensing ☐
   b. Previously fund-holding ☐
   c. Training ☐
   d. Computerised ☐

6. How many general practitioners work in the practice? Full-time ..... Part-time ..... @ .....

7. What is the average length of a routine consultation (booking interval) in your practice? ..... minutes

8. Do you have special clinics for: a. Asthma ☐ b. CHD prevention ☐

9. If computerised, do you have any guidelines for statins prescribing on your computer?
   a. No b. Yes – primary prevention ☐ c. Yes – secondary prevention ☐

10. Do you have a special interest in CHD prevention? a. No ☐ b. Yes ☐

11. Do you prescribe statins for your patients for primary prevention of CHD? a. No ☐
    b. For those at minimum CHD risk of 15% over the next ten years ☐ c. Those at minimum risk of 15-30% ☐
    d. For those at minimum risk of over 30% ☐ e. Yes, but no explicit risk criterion ☐

12. Which version of the Joint British Recommendations guidelines do you use in your practice for secondary prevention? a. None ☐ b. I don't know ☐ c. _____ (specify the publication year)

13. When was the last time that you read or consulted clinical guidelines for CHD prevention?
    a. Never ☐ b. less than a month ago ☐ c. one to three months ago ☐
    d. 3 to 6 month ago ☐ e. 6 to 12 months ago ☐ f. more than 12 months ago ☐

14. Please estimate the number of the patients on your practice list who are eligible for statins for secondary prevention of CHD? a. I don't know ☐ b. _____ (number of patients)

II. YOUR ATTITUDES, PERSPECTIVES AND PLANS ABOUT STATINS PRESCRIBING FOR SECONDARY PREVENTION OF CHD: *Please read questions carefully and tick or draw a circle on whichever applies*

1. Over the past three months, for what proportion of your patients eligible for statins for secondary prevention did you prescribe in accordance with the clinical guidelines?

- ❑ **All patients**
- ❑ **Almost all patients**
- ❑ **More than half the patients**
- ❑ **About half the patients**
- ❑ **Fewer than half the patients**
- ❑ **Very few patients**
- ❑ **No patients**

2. Please estimate how often in the past three months you prescribed statins in your practice in accordance with the clinical guideline recommendations for secondary prevention?

Never 1 2 3 4 5 6 7 **All patients**

3. I intend to follow the clinical guideline recommendations on prescribing statins for secondary prevention over the next three months.

Definitely 1 2 3 4 5 6 7 **Definitely not**

4. How likely is it that you will prescribe statins for secondary prevention of CHD in accordance with the clinical guideline recommendations over the next three months?

Very likely 1 2 3 4 5 6 7 **Very unlikely**

5. I want to prescribe statins for my patients in accordance with the CHD guideline recommendations for secondary prevention over the next three months.

Certainly not 1 2 3 4 5 6 7 **Certainly**

6. For me to prescribe statins for secondary prevention in accordance with the clinical guideline recommendations over the next three months will be

Good	1	2	3	4	5	6	7	**Bad**
Useless	1	2	3	4	5	6	7	**Useful**
Safe	1	2	3	4	5	6	7	**Dangerous**
Appropriate	1	2	3	4	5	6	7	**Inappropriate**
Worthless	1	2	3	4	5	6	7	**Valuable**
Unhelpful	1	2	3	4	5	6	7	**Helpful**

7. If I prescribe statins for secondary prevention in accordance with the clinical guidelines over the next three months, my patients will be healthier

Unlikely 1 2 3 4 5 6 7 **Likely**

* Better health for my patients would be

Extremely bad 1 2 3 4 5 6 7 **Extremely good**

8. If I don't prescribe statins for secondary prevention in accordance with the clinical guidelines over the next three months, I won't get my "quality markers"

Likely 1 2 3 4 5 6 7 **Unlikely**

* Getting "quality markers" would be

Extremely good 1 2 3 4 5 6 7 **Extremely bad**

9. If I don't follow the guideline recommendations on prescribing statins for secondary prevention over the next three months, the standard of care I provide to my patients might be judged to be negligent.

Likely 1 2 3 4 5 6 7 **Unlikely**

* Being judged negligent would be

Extremely bad 1 2 3 4 5 6 7 **Extremely good**

10. Following statin prescribing recommendations in the secondary prevention guidelines over the next three months will prevent harm to my patients.

Unlikely 1 2 3 4 5 6 7 Likely

* Preventing harm to patients would be

Extremely bad 1 2 3 4 5 6 7 Extremely good

11. Following the secondary prevention guideline recommendations on prescribing statins over the next three months will provide a better quality of care for my patients.

Likely 1 2 3 4 5 6 7 Unlikely

* Providing a better quality of care would be

Extremely good 1 2 3 4 5 6 7 Extremely bad

12. People who are important to me think that

I should 1 2 3 4 5 6 7 I should not

follow the secondary prevention guideline recommendations on prescribing statins over the next three months.

13. People whose opinions I value would

approve 1 2 3 4 5 6 7 disapprove

of my prescribing statins in accordance with the secondary prevention guidelines over the next three months.

14. Our practice nurse thinks

I should 1 2 3 4 5 6 7 I should not

follow the secondary prevention guideline recommendations on prescribing statins.

* When prescribing how much do you want to do what your practice nurses think you should?

Not at all 1 2 3 4 5 6 7 Very much

15. My GP colleagues in the practice think

I should 1 2 3 4 5 6 7 I should not

follow the secondary prevention guideline recommendations on prescribing statins.

* When prescribing how much do you want to do what your GP colleagues think you should?

Not at all 1 2 3 4 5 6 7 Very much

16. The local hospital consultants think

I should 1 2 3 4 5 6 7 I should not

follow the guideline recommendations on prescribing statins for secondary prevention.

* When prescribing how much do you want to do what the local hospital consultants think you should?

Not at all 1 2 3 4 5 6 7 Very much

17. The Primary Care Organisation advisers think

I should 1 2 3 4 5 6 7 I should not

follow the secondary prevention guideline recommendations on prescribing statins.

* When prescribing how much do you want to do what the PCT advisors think you should?

Not at all 1 2 3 4 5 6 7 Very much

18. If I wanted to, I could easily prescribe statins for secondary prevention according to the clinical guidelines in my practice over the next three months.

Definitely true 1 2 3 4 5 6 7 Definitely false

19. I don't know if I can follow the secondary prevention guideline recommendations on prescribing statins in my practice over the next three months.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

20. Whether or not I follow the secondary prevention guideline recommendations on the use of statins in my practice over the next three months is entirely up to me

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

21. The guidelines for the use of statins in secondary prevention of CHD are in accordance with evidence.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* Evidence-based clinical guidelines would make following good prescribing practice

More difficult 1 2 3 4 5 6 7 Easier

22. The guidelines for secondary prevention have substantially changed their statin prescribing recommendations.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* Such changes to clinical guidelines would make it

more difficult 1 2 3 4 5 6 7 easier

to prescribe according to their recommendations.

23. I'm always under time pressure to care for my CHD patients.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* Lack of time makes it

more difficult 1 2 3 4 5 6 7 easier

to follow the secondary prevention guideline recommendations on prescribing statins over the next three months.

24. It is appropriate to include the secondary prevention guidelines for the use of statins in the BNF.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* The addition of the clinical guidelines to the BNF would make me

less likely 1 2 3 4 5 6 7 more likely

to following their prescribing recommendations over the next three months.

25. The National Service Framework (NSF) for coronary heart disease advocates the implementation of the secondary prevention guidelines on the use of statins.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* NSFs are powerful influences on prescribing patterns in general practice.

Definitely not 1 2 3 4 5 6 7 Definitely

26. Prescribing statins as recommended by guidelines for secondary prevention could exhaust the prescribing budget of our practice.

Likely 1 2 3 4 5 6 7 Unlikely

* Budgetary limitations makes it

more difficult 1 2 3 4 5 6 7 easier

for me to follow the clinical guidelines' prescribing recommendations over the next three months.

27. Doing an audit of our current practice for secondary prevention of ischaemic heart disease makes it

less likely 1 2 3 4 5 6 7 more likely

that I will follow the guideline recommendations on the use of statins over the next three months.

28. I see myself as an evidence based practitioner.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

29. I see myself as a patient-centred practitioner.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

Many thanks for your valuable time in filling in this questionnaire. Please use the enclosed pre-paid envelope to return the questionnaire to us.

Please use this space for your further comments (continue at the back of the information sheet, if required.)

THE UNIVERSITY of York

Department of Health Sciences

National Survey of General Practitioners
Perceptions and Practice of Asthma Treatment

WE SHALL KEEP ALL RESPONSES **STRICTLY CONFIDENTIAL**

We want to analyse your responses in the context of your prescribing practice. We should therefore be most grateful for permission to use your Prescribing And CosT [PACT) data on asthma and statin prescribing for one year from last quarter. Please sign here if you are willing to give access to those data solely for this research.

1. Do you permit us to access PACT data on your personal prescribing for asthma and of statins? Yes ☐ No ☐
2. Are you the senior partner in the practice? Yes ☐ No ☐
3. If yes, do you permit us to access PACT data for your practice? Yes ☐ No ☐

Signature ..... Date .....

Please fill in the questionnaire even if you do not wish us to access your PACT data.

★★★★

We shall include all returned questionnaires in a lottery.
The winner will receive a **digital camera** early next year.
This is a token of our gratitude for your valued contribution.

I. ABOUT YOU AND YOUR PRACTICE *Please tick whichever applies or write the answer where applicable*

1. Age: ..... 2. Sex: F ☐ M ☐ 3. Year of graduation from medical school? .....

4. Are you:
   Principal in practice Yes ☐ No ☐
   Salaried GP (PMS contract) Yes ☐ No ☐
   Other (specify) .....

5. Practice status: *(please tick all that apply)*
   a. Dispensing ☐
   b. Previously fund-holding ☐
   c. Training ☐
   d. Computerised ☐

6. How many general practitioners work in the practice? Full-time ..... Part-time ..... @ .....

7. What is the average length of a routine consultation (booking interval) in your practice? ..... minutes

8. Do you have special clinics for: a. Asthma ☐ b. CHD prevention ☐

9. If computerised, do you have any guidelines for asthma treatment on your computer?
   a. No ☐ b. Yes ☐

10. Do you have a special interest in asthma? a. No ☐ b. Yes ☐

11. Which version of the BTS asthma guidelines do you use in your practice?
    a. None ☐ b. I don't know ☐ c. _ _ _ _ _ _ _ _ (specify the publication year)

12. When was the last time that you read or consulted the BTS asthma guidelines?
    a. Never ☐ b. less than a month ago ☐ c. one to three months ago ☐
    d. 3 to 6 months ago ☐ e. 6 to 12 months ago ☐ f. more than 12 months ago ☐

13. Please estimate the number of the patients on your practice list who suffer from asthma:
    a. I don't know ☐ b. _ _ _ _ _ _ _ _ (number of patients)

II. YOUR ATTITUDES, PERSPECTIVES AND PLANS ABOUT ASTHMA TREATMENT:

Please read questions carefully and tick or draw a circle on whichever applies

1. Over the past three months, for what proportion of your asthma patients did you prescribe drugs in accordance with the BTS asthma guideline recommendations?

- ❑ **All patients**
- ❑ **Almost all patients**
- ❑ **More than half the patients**
- ❑ **About half the patients**
- ❑ **Fewer than half the patients**
- ❑ **Very few patients**
- ❑ **No patients**

2. Please estimate how often in the past three months you considered the BTS asthma guidelines when prescribing asthmatic drugs in your practice?

Never 1 2 3 4 5 6 7 **Very frequently**

3. I intend to follow the BTS asthma guidelines' prescribing recommendations in my practice over the next three months.

Definitely 1 2 3 4 5 6 7 **Definitely not**

4. How likely is it that you will prescribe drugs for your patients in accordance with the BTS guideline recommendations over the next three months?

Very likely 1 2 3 4 5 6 7 **Very unlikely**

5. I want to prescribe drugs for my patients in accordance with the BTS guideline recommendations over the next three months.

Certainly not 1 2 3 4 5 6 7 **Certainly**

6. For me to prescribe the drugs recommended by BTS asthma guidelines over the next three months will be

Harmful	1	2	3	4	5	6	7	**Beneficial**
Unhelpful	1	2	3	4	5	6	7	**Helpful**
Valuable	1	2	3	4	5	6	7	**Worthless**
Useless	1	2	3	4	5	6	7	**Useful**
Good	1	2	3	4	5	6	7	**Bad**
Appropriate	1	2	3	4	5	6	7	**Inappropriate**

7. If I prescribe in accordance with the BTS asthma guidelines over the next three months, my patients will be healthier.

Unlikely 1 2 3 4 5 6 7 **Likely**

* Better health for my patients would be

Extremely bad 1 2 3 4 5 6 7 **Extremely good**

8. If I don't prescribe in accordance with the BTS asthma guidelines over the next three months, I won't get my "quality markers".

Likely 1 2 3 4 5 6 7 **Unlikely**

* Getting "quality markers" would be

Extremely good 1 2 3 4 5 6 7 **Extremely bad**

9. If I don't follow the BTS prescribing recommendations for asthma over the next three months, the standard of care I provide to my patients might be judged to be negligent.

Likely 1 2 3 4 5 6 7 **Unlikely**

* Being judged negligent would be

Extremely bad 1 2 3 4 5 6 7 **Extremely good**

10. Following prescribing recommendations in the BTS asthma guidelines over the next three months will prevent harm to my patients.

Unlikely **1** **2** **3** **4** **5** **6** **7** **Likely**

* Preventing harm to patients would be

Extremely bad **1** **2** **3** **4** **5** **6** **7** **Extremely good**

11. Following the BTS asthma guidelines' prescribing recommendations over the next three months will provide a better quality of care for my patients

Likely **1** **2** **3** **4** **5** **6** **7** **Unlikely**

* Providing a better quality of care would be

Extremely good **1** **2** **3** **4** **5** **6** **7** **Extremely bad**

12. I am expected to follow the BTS prescribing recommendations for my asthmatic patients over the next three months.

True **1** **2** **3** **4** **5** **6** **7** **False**

13. People whose opinions I value would

approve **1** **2** **3** **4** **5** **6** **7** **disapprove**

of my prescribing in accordance with the BTS asthma guidelines over the next three months.

14. Our practice nurse thinks

I should **1** **2** **3** **4** **5** **6** **7** **I should not**

follow the BTS asthma guidelines' prescribing recommendations.

* When prescribing how much do you want to do what your practice nurses think you should?

Not at all **1** **2** **3** **4** **5** **6** **7** **Very much**

15. My GP colleagues in the practice think

I should **1** **2** **3** **4** **5** **6** **7** **I should not**

follow the BTS asthma guidelines' prescribing recommendations.

* When prescribing how much do you want to do what your GP colleagues think you should?

Not at all **1** **2** **3** **4** **5** **6** **7** **Very much**

16. The local hospital consultants think

I should **1** **2** **3** **4** **5** **6** **7** **I should not**

follow the BTS asthma guidelines' prescribing recommendations.

* When prescribing how much do you want to do what the local hospital consultants think you should?

Not at all **1** **2** **3** **4** **5** **6** **7** **Very much**

17. The Primary Care Organisation advisers think

I should **1** **2** **3** **4** **5** **6** **7** **I should not**

follow the BTS asthma guidelines' prescribing recommendations.

* When prescribing how much do you want to do what the PCT advisors think you should?

Not at all **1** **2** **3** **4** **5** **6** **7** **Very much**

18. For me to prescribe medicines for asthmatic patients according to the BTS guideline recommendations over the next three months would be

Difficult **1** **2** **3** **4** **5** **6** **7** **Easy**

19. I don't know if I can follow the BTS asthma guidelines' prescribing recommendations in my practice over the next three months.

Strongly agree **1** **2** **3** **4** **5** **6** **7** **Strongly disagree**

20. How much control do you feel you have over prescribing medicines in accordance with the BTS asthma guidelines over the next three months?

No control 1 2 3 4 5 6 7 Complete control

21. The BTS asthma guidelines are in accordance with evidence.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* Evidence-based clinical guidelines would make good prescribing

More difficult 1 2 3 4 5 6 7 Easier

22. The BTS asthma guidelines have substantially changed their prescribing recommendations.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* Such changes to clinical guidelines would make it

more difficult 1 2 3 4 5 6 7 easier

to prescribe according to their recommendations.

23. I'm always under time pressure to care for my asthmatic patients.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

* Lack of time makes it

more difficult 1 2 3 4 5 6 7 easier

for me to follow the BTS asthma guidelines prescribing recommendations over the next three months.

24. The BTS asthma guidelines are not flexible enough to help an individual patient.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* Limited options within the clinical guidelines make my prescribing for individual patients

less likely 1 2 3 4 5 6 7 more likely

to accord with their recommendations over the next three months.

25. It is appropriate to include the BTS asthma guidelines in the British National Formulary (BNF).

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* The addition of the clinical guidelines to the BNF would make me

less likely 1 2 3 4 5 6 7 more likely

to follow their prescribing recommendations over the next three months.

26. I continually hear about the BTS asthma guidelines.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

* Hearing about the clinical guidelines from different sources makes me

less likely 1 2 3 4 5 6 7 more likely

to follow their prescribing recommendations.

27. I see myself as an evidence based practitioner.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

28. I see myself as a patient-centred practitioner.

Strongly agree 1 2 3 4 5 6 7 Strongly disagree

Many thanks for your valuable time in filling in this questionnaire. Please use the enclosed pre-paid envelope to return the questionnaire to us.

Please use this space for your further comments (continue at the back of the information sheet, if required.)

Questions and answers about this study

What is the background to this survey?

Clinical guidelines are increasingly used to influence the behaviour of health professionals in primary care. However, previous studies have revealed that general practitioners have some concerns about the clinical guidelines. This national survey aims to inform the future development and use of clinical guidelines in primary care in the UK. It forms a major part of the project we call **Studying Adherence to Guidelines and Evidence (SAGE)**.

Who are the researchers?

The researchers are independent, university-based academics. Arash Rashidian is a medical practitioner and Health Services Researcher at University of York Department of Health Sciences. Ian Russell is a Fellow of Royal College of General Practitioners and Founding Professor of Health Sciences in the same department.

How was this questionnaire developed?

We developed the questionnaire with the generous help of several GPs who took part in a qualitative study and others who piloted the questionnaire. It focuses mainly on prescribing statins for secondary prevention of coronary heart disease and the influence of clinical guidelines. We used the Theory of Planned Behaviour as a base for the questionnaire.

What is the Theory of Planned Behaviour?

This theory has been successfully used to explain many health related behaviours of patients. However, few researchers have used it to explain health professionals' behaviour. We believe this is the first use of the TPB to study the influence of clinical guidelines on prescribing behaviour of GPs. The advantage of this approach is its ability to reveal GPs' attitudes towards, and concerns about, clinical guidelines in a rigorous way.

How many doctors are being studied?

In this survey 1000 GPs from 50 Primary Care Organisations (PCOs) in England will receive the questionnaires. They are a random sample of all GPs working in England, stratified by the demographic characteristics of PCOs in which they work.

Do I have to complete the questionnaire?

No, the choice is yours. But we hope you will help. If you are uncertain about some of the questions, please try your best to choose the most appropriate answer. If you are unwilling to take part, please say so and we will not send you any reminders.

Will I be paid for filling in the questionnaire?

Unfortunately not. Instead we shall include all returned questionnaires in a lottery. The winner will receive a **digital camera** early next year. This is a token of our gratitude for your valued contribution.

What is the number on the questionnaire for?

We shall use these numbers to log the questionnaires as they return. This will prevent us from sending reminders to GPs who have replied.

Why do you need my PACT data?

These data will enable us to analyse your questionnaire in the context of your prescribing practice. They will be used solely for this purpose. Please fill in the questionnaire even if you do not wish us to access your PACT data.

How shall you protect my patients' confidentiality?

PACT data can not be linked to the individual patients.

How shall you protect my confidentiality?

All information about individuals or practices will be held strictly confidential and will not be disclosed or released to others outside the research team. All reports and publications will aggregate data to prevent identification of individuals or practices.

How the results will be used?

The results will be disseminated through presentation at relevant conferences and submitting to peer reviewed international journals. We hope they will improve the use of clinical guidelines in future.

Please use this space for your further comments, if required, and return it with the questionnaire.

THE UNIVERSITY of York

Studying Adherence to Guidelines and Evidence (SAGE)

Arash Rashidian, MD
Health Services Research Scholar
Ian Russell, PhD, FRCGP, FRCP Edin
Founding Professor of Health Sciences

DEPARTMENT OF
HEALTH SCIENCES
Donald Irvine Wing
Alcuin College
Heslington, York YO10 5DD
Telephone (01904) 434498
Fax (01904) 434517
Email ar130@york.ac.uk

22 January 2002

Dr «GPINITIALS» «GPSURNAME»
«Add1»
«Add2»
«Add3»
«Add4»
«Postcode»

Dear Dr «GPSURNAME»

National Survey of Statin Prescribing by General Practitioners

We should be most grateful for your help with this survey. **Our primary aim is to help general practice and general practitioners.** We are keen to improve the development and use of clinical guidelines in primary care in the UK.

We know the profession is under increasing pressure. So we should be especially grateful if you would give us (and thus, we hope, your profession!) a little of your valuable time to complete the enclosed questionnaire. The questionnaire focuses mainly on statin prescribing for secondary prevention of CHD. With your co-operation we believe that it can improve understanding of GPs' attitudes to, and concerns about, guidelines. So please take **ten minutes** to fill in the questionnaire and use the enclosed business reply envelope to send it back to us.

We should also be most grateful for access to your Prescribing And CosT (PACT) data on asthma and statin prescribing for one year from last quarter. This will enable us to analyse your questionnaire in the context of your prescribing practice. We shall take great care with this analysis. In particular we shall keep **strictly confidential** all information that would permit identification of an individual or practice.

As a token of our gratitude we shall include all returned questionnaires in a lottery. The winner will receive a **digital camera** early this year.

Many thanks and kind regards
Yours sincerely

Prof Ian Russell FRCGP

Dr Arash Rashidian MD

Appendix VIII. Research proposal: A theory-based approach for promoting uptake of clinical guidelines: a case-study of osteoporosis

Summary

Introduction Clinical guidelines are increasingly used to promote improved general practitioners (GPs) prescribing, however GPs' adherence varies widely. An approach based on psychological theories such as the TPB may enhance understanding of this variation and identify more effective methods of changing behaviour. However the utility of this theory in explaining and changing doctors' behaviour has not been rigorously evaluated.

Aim This study will explore and test the ability of a psychological model of behaviour to explain, predict and improve GPs' adherence to national clinical guidelines.

Main objectives

1. To explain and predict variation in GPs' prescribing intentions and prescribing behaviour of drugs covered in the forthcoming National Institute of Clinical Excellence (NICE) clinical guideline for osteoporosis using TPB.
2. To evaluate the effect of a behavioural intervention on prescribing intentions and behaviour in accordance with this NICE clinical guideline.
3. To evaluate the cost-effectiveness of this intervention in changing GPs' prescribing.

Design and setting A randomised controlled trial in English and Welsh general practice.

Participants and methods A small qualitative study of GPs will be used to identify the salient beliefs about prescribing for osteoporosis. These will be used to develop and pilot a TPB questionnaire to elicit GPs' beliefs, attitudes, subjective norms and perceived behavioural controls and other factors related to the osteoporosis prescribing. We shall explore their stated osteoporosis prescribing intentions and behaviour using vignettes. The questionnaire and prescribing data will be used to obtain the baseline data of a random sample of 500 GPs. We shall use the results to develop an appropriate behavioural intervention to improve GPs' uptake of the NICE osteoporosis guideline. We shall evaluate this intervention through stratified random allocation of participating GPs to control (receiving NICE guideline) and intervention (receiving NICE guideline and intervention) groups. GPs will be followed up using both repeating the questionnaire and analysis of their prescribing data. We shall expect to obtain complete data for 400 GPs. The analysis will a) assess the utility of TPB in explaining, predicting and improving GPs' prescribing, b) estimate the incremental improvement in osteoporosis prescribing as a result of this intervention, and c) estimate costs of the intervention.

Full proposal:

1. Title

A theory-based approach for promoting uptake of clinical guidelines: a case-study of osteoporosis

2. Purpose

Study aim This study will explore and test the ability of a psychological model of behaviour change to explain, predict and improve GPs' adherence to national clinical guidelines.

Objectives

1. To use the TPB to explain and predict variation in GPs' prescribing intentions and prescribing behaviour of drugs covered in the forthcoming National Institute of Clinical Excellence (NICE) clinical guideline for osteoporosis.
2. To evaluate the effect of a behavioural intervention based on the TPB, on prescribing intentions and behaviour in accordance with this NICE clinical guideline.
3. To evaluate the cost-effectiveness of this behavioural intervention in changing prescribing.

This research project will incorporate five of health services research disciplines: health psychology, epidemiology, health economics, biostatistics and health policy. It has the potential to develop a theoretically sound evidence-based approach which can be used by primary care organisations for future implementation strategies and by guideline producing organisations to improve the uptake of guidelines.

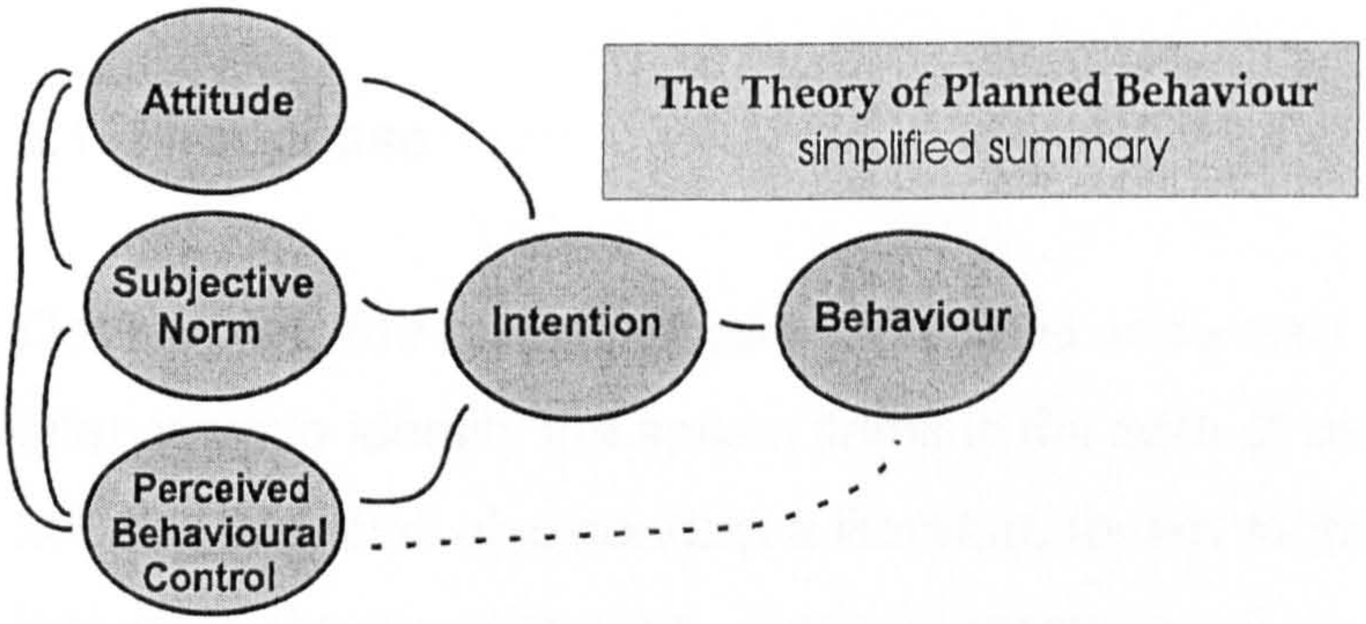
3. Background

Clinical guidelines and change in behaviour Recent reviews(NHS Centre for Reviews and Dissemination, 1999) have shown that changing professional behaviour is difficult (Bero et al, 1998; NHS Centre for Reviews and Dissemination, 1999). Clinical guidelines are increasingly used to try to

improve prescribing by GPs, however GPs vary widely in their adherence to clinical guidelines. Nevertheless some interventions are more successful than others in improving their uptake (NHS Centre for Reviews and Dissemination, 1994; Wensing et al, 1998).

Theoretical approach to prescribing behaviour change Much of clinical behaviour change literature is based on the naïve assumption that clinicians will change if they are given information (information deficit model) and few studies have tested the theoretical models that acknowledge the psychological and organisational processes which precede behaviour (Grol, 1997; Marteau et al, 2002). Using theory informed approaches might enhance the effectiveness of guideline implementation in general and appropriate prescribing in particular (Raisch, 1990b; Grol, 1997; Rashidian and Russell, 2003). TPB(Conner and Sparks, 1996) states that "individuals make behavioural decisions based upon consideration of available information" (Ajzen, 1991; Conner and Sparks, 1996). According to TPB human behaviour is guided by three kinds of considerations: beliefs about the likely consequences of the behaviour, beliefs about the normative expectations of others, and beliefs about the presence of factors that may facilitate or impede (control) the behaviour (Ajzen, 1991; Ajzen, 2002b). These three types of beliefs produce 'attitudes', perceived social pressure ('subjective norms') and 'perceived behavioural controls', respectively. These in turn will form 'intentions' and predict actual behaviour (Ajzen, 1991; Conner and Sparks, 1996; Ajzen, 2002b). Some evidence suggests that TPB might explain variation in health professionals' intentions (including GPs' prescribing intentions) and self-reported behaviours (Millstein, 1996; Lambert et al, 1997; Levin, 1999; Walker et al, 2001; Puffer and Rashidian, 2004). The applicant has performed two cross-sectional surveys to test the utility of TPB in explaining GPs prescribing (data entry and analysis in progress). The Medical Research Council recently funded a project to compare utility of six different psychological theories in predicting GPs and dental practitioners behaviour using seven cross-sectional surveys (Eccles et al, 2001a). Although there are several studies which indicate this theory may have potential in explaining variance (Armitage and Conner, 2001) and developing behaviour change interventions (Hardeman et al, 2002), the utility of TPB in explaining and changing health professionals' behaviour has not been rigorously evaluated. This study has the potential to

improve our understanding of theories of health professional behaviour change and their value as the basis of interventions for changing prescribing behaviour in a few fronts.



Osteoporosis as a tracer condition Prevention and treatment of osteoporosis is among the NHS priorities (Department of Health, 2001a). Up to 14,000 people a year die in the UK as a result of an osteoporotic hip fracture and 1 in 3 women and 1 in 12 men over 50 are affected by osteoporosis and almost half of all women experience an osteoporotic fracture by the time they reach the age of 70 (Department of Health, 2001a). The estimated social and acute care costs of hip fractures are more than £1.7 billion annually (Dolan and Torgerson, 1998; Torgerson et al, 2001). Osteoporosis is often missed by patients and GPs in their consultations. It is most prevalent in post-menopausal women who may also suffer from other age related diseases. GPs' potential conflicting treatment intentions (i.e. treating more than one disease or risk factor), short period of consultation and patients' own agenda all suggest that psychological processes might play an important role in osteoporosis prevention and treatment. The introduction by NICE of an osteoporosis clinical guideline is also timely for this application and our proposed research plan (see section 4.5. - project milestones).

4. Study design and methods

This study is a randomised controlled trial in English and Welsh general practice.

Participants We shall recruit GPs who are practicing in North & East Yorkshire & North Lincolnshire (the catchment area of the Hull York Medical School) and North & West of Wales (the catchment areas of the North Wales & Swansea Clinical Schools) and whose practice list size is larger than 1000.

4.1. First phase

Developing the questionnaire We shall undertake 10-15 semi-structured interviews to identify the salient items in the care of osteoporosis (Ajzen, 2002b). We shall also conduct a literature review to identify other variables to include in the questionnaire. Each interviewee will be offered £50 for participating. We shall then pilot the questionnaire on a judgemental sample of 40 GPs. The questionnaire will include questions about the GP, practice organisational characteristics and psychological items and will take about 15 minutes to complete. Psychological items will include osteoporosis prescribing specific intentions, attitudes, subjective norms, and perceived behavioural controls and their relevant beliefs (Ajzen, 1991; Conner and Sparks, 1996; Gollwitzer, 1999). We shall measure stated behaviour based on responses to vignettes. Through these interviews and the literature review we shall seek to improve the predictive utility of TPB by inclusion of other psychological items relevant to osteoporosis prescribing (Conner and Armitage, 1998; Terry et al, 1999).

First survey We shall approach 1000 GPs. All GPs will be sent the questionnaire, a covering letter, an information sheet and a business reply envelope. We shall ask of GPs consent for access to their osteoporosis related prescribing data for a period of two years and the intervention, and inform them of our planned second survey. We anticipate this will yield 500 completed questionnaires.

4.2. Second phase

Intervention We shall use the results of the first survey to develop an appropriate behavioural intervention (Ajzen, 2002a) to improve GPs' uptake of NICE osteoporosis guideline. Those 500 GPs who have responded to the first

survey will be stratified and randomised to control (receiving NICE guideline) and intervention (receiving NICE guideline and intervention) groups. We shall use practice list size and a TPB indicator variable for stratification. The intervention will include the delivery by post of educational leaflets and other items, based on TPB in order to promote the uptake of the osteoporosis guideline. In collaboration with NICE, we shall arrange to have access to the earlier drafts of the guideline before its final publication to prepare the intervention, so that it can be undertaken soon after the introduction of the guideline.

Second survey GPs will be followed up by repeating the questionnaire and analysis of their prescribing data. We anticipate 400 of the 500 GPs responding to the questionnaire.

4.3. Proposed sample size and response rate We aim to recruit 400 GPs to complete both of the first and second surveys. This way we shall have 80% power at 5% significance level to detect a mean change of 0.1 in osteoporosis prescribing rate (prescribed item per weighted practice list size) at a standard deviation of 0.35. Following Green suggestions we shall require a minimum sample of 50+8m=130 (m=10 is number of predictor variables in the regression equation) for testing the multiple correlation (Green, 1991; Tabachnick and Fidell, 2001). We have also previously suggested that a sample size of 191 is adequate for a TPB survey of health professionals (Rashidian et al, 2002). In recent years the response rate has been falling within general practice (Kaner et al, 1998), and some surveys achieve less than the average response rate of 54% observed in 'published studies' (Asch et al, 1997). Considering different strategies to improve response rate (McColl et al, 2001) (including sending two reminders for each survey), we shall aim to achieve a 50% response rate in our first survey and 80% when we survey the respondents for the second time. Therefore we will approach 1000 GPs to receive 500 completed questionnaires in first survey and 400 in second one (200 per group). For every returned questionnaire we shall offer £20.

4.4. Outcomes and analysis Prescribing outcomes are change in prescribed items and their costs (numerators) from weighted practice list size

(denominator), in the year before and after the intervention. Practice list size will be weighted using age-sex specific categories (e.g. restricting to women over 50). We shall obtain osteoporosis related prescribing recorded in Prescriptions And CosT (PACT) data as a (proxy) measure of prescribing behaviour. Those include calcium, Vitamin D, selective oestrogen receptor modulators, bisphosphonates and parathyroid hormone (subject to licensing). Choice of target medicines depends on the clinical guideline recommendations. We shall measure prescribing at the practice level and will use a practice level deprivation index to adjust the findings (Lloyd et al, 1995b). To assess the utility of TPB in explaining and predicting GPs' prescribing behaviour in accordance with the clinical guideline, we shall use multiple regression to regress proxy measure of prescribing behaviour on indicator variables within the TPB-based model measured by questionnaire. We shall use regression analysis and structural equation modelling (Baron and Kenny, 1986; Hankins et al, 2000; Tabachnick and Fidell, 2001), using SPSS and LISREL software packages (Hayduk, 1987). We shall perform an economic evaluation on our intervention using cost-effectiveness analysis. Costs are the incremental costs of the intervention. Effectiveness will be measured in terms of improvement in prescribing behaviour and also intention to prescribe in accordance with the guideline. This is particularly important as it is unlikely to be cost-effective to seek to achieve complete compliance with clinical guidelines (Mason et al, 2001). We shall use sensitivity analysis and other modelling techniques to address uncertainties in the data and to estimate the potential improvement in patients health as a result of observed change in prescribing (Buxton et al, 1997; Brennan and Akehurst, 2000; Weinstein et al, 2001). Owing to the clustered nature of our data, we shall perform multi-level analysis (Rice and Leyland, 1996; Rice and Jones, 1997) and allow for panel nature of our surveys (Tabachnick and Fidell, 2001).

4.5. Project milestones

Dates	Months	Stage
Oct 03 – Jan 04	4	Establishing contact with guideline developers. Review of relevant osteoporosis literature. Negotiating populations for the main survey.
Feb – May 04	4	Semi-structured interviews and developing questionnaire for first survey of GPs.
Jun – Sep 04	4	Pilot survey and revision of questionnaire for first survey.
Oct – Nov 04	2	First survey, 1st reminder after 20 days and 2nd reminder after 40 days.
Dec 04 – May 05	6	Data entry and preliminary analysis of first survey. Randomisation of participants to control and intervention groups. Preparation of educational package.
Jun 05	1	Distribution of educational package to intervention group and simpler related material to control group.
Jul – Oct 05	4	Start writing-up first survey. Refine prescribing and cost data. Plan for economic modelling and sensitivity analysis.
Nov – Dec 05	2	Second survey, 1st reminder after 20 days and 2nd reminder after 40 days.
Jan – Apr 06	4	Data entry and preliminary analysis of second survey. Collect data for economic modelling and sensitivity analysis.
May – Jun 06	2	Final refining and preliminary analysis of prescribing data.
Jul – Sep 06	3	Final data analysis and writing up.

5. Research team This protocol has been developed in collaboration with the applicant's proposed supervisors (Prof Ian Russell and Prof Ian Watt), Head of York Department of Health Sciences (Prof Trevor Sheldon, on methodology and planning) and his advisers and collaborators who are Dr David Torgerson (York, on economic evaluation and osteoporosis), Prof Marie Johnston (St Andrews, on psychological theories of behaviour change), Dr Nigel Rice (York, on multi-level analysis) and Prof Martin Eccles (Newcastle, on clinical guideline implementation).

6. Ethics approval We shall approach one of the Northern & Yorkshire or Wales Multi-centre Research Ethics Committees (MREC) in May 2003. This is due to the fact that MRECs no longer accept applications for ethical review unless funding has been agreed. We shall then apply to all relevant Local Research Ethics Committees for locality assessment only.

Abbreviations

ACE	Angiotensin Converting Enzyme
ADQ	Adequate Daily Quantity
AIDS	Acquired Immunodeficiency Syndrome
ASTRO-PUs	Age, Sex, and Temporary Resident Originated Prescribing Units
ATT	Attitude
BA	Before-After (study design)
BEH	Behaviour
BI	Behavioural Intention
BIB	Balanced Incomplete Block (study design)
BMA	British Medical Association
BMJ	British Medical Journal
BNF	British National Formulary
BTS	British Thoracic Society
CBA	Controlled Before-After trial
CCT	Controlled Clinical Trial
CDSS	Clinical Decision Support Systems
CME	Continuous Medical Education
CHD	Coronary Heart Disease
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
CPD	Continuous Professional Development
CPG	Clinical Practice Guideline
CQI	Continuous Quality Improvement
CRD	Centre for Reviews and Dissemination (University of York, UK)
CRCT	Cluster-Randomised Controlled Trial
DARE	Database of Abstracts of Reviews of Effects
DIN	Doctors Independent Network
DDD	Defined Daily Dose
DUE	Drug Use Evaluation

DUR	Drug Utilisation Review
EBM	Evidence-based Medicine
EBHC	Evidence-based Health Care
EPOC	Effective Practice and Organisation of Care (Cochrane Review Group)
ESRC	Economic & Social Research Council
FDA	Food and Drug Administration (USA body)
GMS	General Medical Services
GP	General Practitioner
GPRD	General Practice Research Database
HDL	High-Density Lipoprotein
HIV	Human Immunodeficiency Virus
HMGCoA	3-hydroxy-3-methylglutaryl coenzyme A
HMO	Health Maintenance Organisation
HTA	Health Technology Assessment
ICC	Intra-Cluster Coefficient
IP	Inter-professional
ISTAHC	the International Society for Technology Assessment in Health Care
JBR	Joint British Recommendations
LDL	Low-Density Lipoprotein
LSHTM	the London School of Hygiene & Tropical Medicine
MI	Myocardial Infarction
MRC	Medical Research Council
N/A	Not Applicable
N/R	Not Reported
NHS	National Health Service
NIC	Net Ingredient Cost
NICE	the National Institute for Clinical Excellence (UK NHS Special Health Authority)
NS	Non Significant
NSAID	Non-Steroidal Anti-Inflammatory Drug
NSF	National Service Framework

OLS	Ordinary Least Square
ORS	Oral Re-hydration Solution
OTC	Over The Counter
PACT	Prescribing Analyses and CosT
PBC	Perceived Behavioural Control
PCG	Primary Care Group
PCO	Primary Care Organisation
PCP	Primary Care Physician
PCT	Primary Care Trust
PMS	Personal Medical Services
PPA	the Prescription Pricing Authority (UK NHS Special Health Authority)
PPI	Proton Pump Inhibitor
PRODIGY	Prescribing Rationally with Decision support in General Practice (computer software)
PSA	Prostate Specific Antigen
PSU	the Prescribing Support Unit (UK body)
PU	Prescribing Unit
RCT	Randomised Controlled Trial
R&D	Research and Development
SAGE	Study of Adherence to Guidelines and Evidence
SD	Standard Deviation
SE	Standard Error
SEM	Structural Equation Modelling
SIGN	the Scottish Intercollegiate Guideline Network
SN	Subjective Norm
SSRI	Selective Serotonin Reuptake Inhibitor
STAR-PUs	Specific therapeutic group age-sex related prescribing units
STD	Sexually Transmitted Disease
STG	Standard Treatment Guideline
TPB	the Theory of Planned Behaviour
TQM	Total Quality Management

TRA	the Theory of Reasoned Action
UK	the United Kingdom
USA	the United States of America
UTI	Urinary Tract Infection
VIF	Variance Inflation Factor
WHO	World Health Organisation

Glossary

Academic detailing	See 'educational outreach'
Adequate Daily Quantity	ADQ. It is the British equivalent of DDD. An analytical unit produced in order to compare more accurately the prescribing activity of primary care practitioners (Prescribing Support Unit, 2002)
Adoption	Health care providers' commitment and decision to change their practices; the actual change in practice (Davis and Taylor-Vaisey, 1997)
Audit and feedback	Any summary of clinical performance of health care over a specified period of time. The summary may also include recommendations for clinical action. The information may be given in a written, verbal or electronic format (Thomson O'Brien et al, 1997)
Clinical decision support systems	See 'computer-based CDSS' and 'medical decision aids'
Clinical guidelines	See 'clinical practice guideline'
Clinical practice guidelines	Systematically developed statements to assist decisions about appropriate care for specific clinical circumstances (Field and Lohr, 1990) p 38. See also 'medical review criteria' and 'standard of quality'
Computer-based CDSS	Computer software using a knowledge base designed for use by a clinician involved in patient care as a direct aid to clinical decision making. Characteristics of an individual patient are matched to information in the knowledge base. Patient-specific information in the form of assessments or recommendations is presented to the clinician (Johnston et al, 1994). See also 'medical decision aids'
Conferences	See 'educational meetings' (Freemantle et al, 1997)
Continuous quality improvement	A philosophy of continual improvement of the processes associated with providing a good or service that meets or exceeds consumer expectations (Shortell et al, 1998)
Defined Daily Doses	DDD. A measure of the amount of drug prescribed based on standard therapeutic units. The number of defined daily doses is calculated by dividing the

	total amount of a drug prescribed by the defined daily dose for the drug (Majeed et al, 1997) OR the assumed average maintenance dose per day for a drug used on its main indication in adults (WHO Collaborating Centre for Drug Statistics Methodology, 1999)
Diffusion	Distribution of information and practitioners' natural, unaided adoption of policies and practices (Davis and Taylor-Vaisey, 1997)
Dissemination	Communication of information to clinicians to improve their knowledge or skills; more active than diffusion, dissemination targets a specific clinical audience (Davis and Taylor-Vaisey, 1997)
Educational materials	Distribution of published or printed recommendations for clinical care, including clinical practice guidelines, audio-visual materials and electronic publications (Jamtvedt et al, 2003)
Educational meetings	Participation of health care providers in conferences, lectures, workshops or traineeships outside the providers' practice settings (Jamtvedt et al, 2003)
Educational outreach	Short, one-to-one conversations between a detailer and a practitioner, with a goal of persuading the detailee to change behaviour through useful information and evidence (Gross and Pujat, 2001) OR Use of a trained person who meets with providers in their practice settings to provide information with the intent of changing the provider's performance. The information given may include feedback about performance (Thomson O'Brien et al, 1997)
Educationally influentials	See 'local opinion leaders' (Lomas and Haynes, 1988)
Evidence-based health care	See 'evidence-based medicine'
Evidence-based medicine	Evidence-based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research (Sackett et al, 1996) (p 71)
Expert systems	See 'computer-based CDSS' and 'medical decision aids'
General practice	See 'general practitioner'

General practitioner	The general practitioner is a specialist trained to work in the front line of a healthcare system and to take the initial steps to provide care for any health problem(s) that patients may have. The general practitioner takes care of individuals in a society, irrespective of the patient's type of disease or other personal and social characteristics, and organises the resources available in the healthcare system to the best advantage of the patients. The general practitioner engages with autonomous individuals across the fields of prevention, diagnosis, cure, care, and palliation, using and integrating the sciences of biomedicine, medical psychology, and medical sociology (Olesen et al, 2000, p355).
Health services research	Health services research is the multidisciplinary field of scientific investigation that studies how social factors, financing systems, organisational structures and processes, health technologies, and personal behaviours affect access to health care, the quality and cost of health care, and ultimately our health and well-being. Its research domains are individuals, families, organisations, institutions, communities, and populations (Lohr and Steinwachs, 2002)
Implementation	Putting a guideline in place; more active than dissemination, it involves effective communication strategies and identifies and overcomes barriers to change by using administrative and educational techniques that are effective in the practice setting (Davis and Taylor-Vaisey, 1997)
Individual instruction	See 'educational outreach' (Grol, 1992; Wensing and Grol, 1994)
Knowledge translation	The exchange, synthesis and ethically sound application of knowledge—within a complex system of interactions among researchers and users—to accelerate the capture of the benefits of research . . . through improved health, more effective services and products, and a strengthened health care system (Davis et al, 2003)
Local consensus processes	Inclusion of participating providers in discussion to ensure that they agreed that the chosen clinical problem was important and the approach to managing the problem was appropriate (Jamtvedt et al, 2003)
Local opinion leader	Use of providers nominated by their colleagues as 'educationally influential'. The investigators must explicitly state that the opinion leaders were identified by their colleagues (Freemantle et al, 1997)

Marketing	See 'tailoring' (Freemantle et al, 1997)
Medical decision aids	Active knowledge systems which use two or more items of patient data to generate case specific advice (Johnston et al, 1994). See also 'computer-based CDSS'
Medical review criteria	Systematically developed statements that can be used to assess the appropriateness of specific health care decisions, services, and outcomes (Field and Lohr, 1990) p 44
Net Ingredient Cost	NIC. Cost of the drug before discounts excluding any dispensing costs or fees. It also excludes any adjustment for income obtained where a prescription charge is paid at the time the prescription is dispensed or where the patient has purchased a pre-payment certificate
NICE	National Institute for Clinical Excellence. A special NHS authority involved in introduction new guidelines and guidance for the NHS.
Outreach visits	See 'educational outreach'
PACT data	Prescribing Analyses and Cost. Information on GPs' prescribing obtained from prescriptions dispensed by community pharmacists, dispensing GPs, and appliance contractors (Majeed et al, 1997)
Patient-centred care	A philosophy of care that encourages: (a) shared control of the consultation, decisions about interventions or management of the health problems with the patient, and/or (b) a focus in the consultation on the patient as a whole person who has individual preferences situated within social contexts (in contrast to a focus in the consultation on a body part or disease) (Lewin et al, 2001)
Patient mediated interventions	Any intervention aimed at changing the performance of health care providers indirectly by providing information, prompts, or support to the patient (Gill et al, 1999; Jamtvedt et al, 2003)
PPA	Prescription Prescribing Authority. A special NHS authority located in Newcastle, England. The PPA is the source of PACT data
Primary care physician	Primary care physicians are medically qualified physicians who provide primary health care. Primary health care provides integrated, easy to access, health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained and continuous relationship with

	patients, and practising in the context of family and community (Vanselow et al, 1995) OR Medical health care professionals providing first contact and on-going care to patients, regardless of the patient's age, gender or presenting problem (Bower and Sibbald, 1999)
Prescribing support unit	PSU. It is located in Leeds and funded by the Department of Health for conducting research on, and analysis of UK prescribing data.
Protocol	A comprehensive set of criteria for a single clinical condition or aspects of organisation (Baker and Fraser, 1995) (p 370). For 'criteria' see 'medical review criteria'
Prescribing unit	PU. A measure of patients' needs for prescribed drugs weighted for age; patients aged under 65 years count as one unit, patients aged 65 and over count as three units (Majeed et al, 1997)
Providers	Health care professionals, including physicians
Public interest detailing	See 'educational outreach' (Thomson O'Brien et al, 1997)
Quality of care	Quality of care is the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge (Lohr and Institute of Medicine, 1990) (p 21)
Rational use of drugs	The rational use of drugs requires that patients receive medications appropriate to their clinical needs, in doses that meet their own individual requirements for an adequate period of time, and at the lowest cost to them and their community (Le Grand et al, 1999) from 'WHO conference of experts Nairobi 1985'
Reminders	Any intervention, manual or computerised, that prompts the health care provider to perform a clinical action (Jamtvedt et al, 2003)
Research utilisation	A process directed toward transfer of specific research-based knowledge into practice through systematic use of a series of activities (Logan and Graham, 1998)
Review criteria	See 'medical review criteria'
Setting	The practice site – not so much its location, although this may be important, as its type – the setting may also imply, but not define, aspects of workload, relevant health care team members,

	mix of patients and funding mechanisms (Davis and Taylor-Vaisey, 1997)
Standard	See 'standard of quality'. Also: the percentage of events that should comply with the criterion (Baker and Fraser, 1995) p 370. For 'criterion' see 'medical review criteria'
Standard treatment guidelines	See 'clinical practice guidelines'. This term is commonly used in the literature from developing countries instead of clinical practice guidelines.
Standard of quality	Authoritative statement of minimum levels of acceptable performance or results, excellent levels of performance or results, or the range of acceptable performance or results (Field and Lohr, 1990) p 50
Tailoring	Use of personal interviewing, group discussion ('focus groups'), or a survey of targeted providers to identify barriers to change and subsequent design of an intervention that addresses identified barriers (Jamtvedt et al, 2003)

References

Abraham C, Clift S, and Grabowski P (1999). Cognitive predictors of adherence to malaria prophylaxis regimens on return from a malarious region: a prospective study. *Social Science & Medicine,* 48, 1641-1654.

Abraham C, Sheeran P, and Johnston M (1998). From health beliefs to self-regulations: theoretical advances in the psychology of action control. *Psychology and Health,* 13, 569-591.

Ackermann SP and Cheal N (1994). Factors affecting physician adherence to breast cancer screening guidelines. *Journal of Cancer Education,* 9, 96-100.

Adler M and Ziglio E (1996). *Gazing into the oracle: the Delphi method and its application to social policy and public health.* London and Bristol: Jessica Kingsley Publishers.

Agnew CR (2000). Behavioural and normative beliefs about condom use: comparing measurement alternatives within the theory of reasoned action. In *Understanding and changing health behaviour: from health beliefs to self-regulation.* Norman P, Abraham C, and Conner M (eds). pp 115-135. Amsterdam: Harwood Academic Publishers.

Ajzen I (1985). From intentions to actions: the theory of planned behavior. In *Action-control: from cognition to behavior.* Kuhl J and Beckmann J (eds). pp 11-39. Berlin and New York: Springer-Verlag.

Ajzen I (1988). *Attitudes, personality and behavior.* Dorsey Press: Chicago.

Ajzen I (1991). The theory of planned behavior. *Organisational Behavior and Human Decision Processes,* 50, 179-211.

Ajzen I (1998). Models of human social behaviour and their application. *Psychology and Health,* 13, 735-740.

Ajzen I (2002a). Behavioral interventions based on the theory of planned behavior

Ajzen I (2002b). Construction of a standard questionnaire for the theory of planned behavior *http://www-unix oit umass edu/~aizen/,*

Ajzen I and Fishbein M (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice-Hall.

Allery LA, Owen PA, and Robling MR (1997). Why general practitioners and consultants change their clinical practice: a critical incident study. *BMJ,* 314, 870-874.

Altman DG (1991). *Practical statistics for medical research.* London: Chapman & Hall.

Amidi S, Solter S, Rashidian B, Zokaian AR, and Razmjoian F (1975). Antibiotic use and abuse among physicians in private practice in Shiraz, Iran. *Medical Care,* 13, 341-345.

Amis M (2000). *All from experience.* London: Jonathan Cape.

Anderson AS, Cox DN, McKellar S, Renolds J, Lean M, and Mela D (1998). Take Five, a nutrition education intervention to increase fruit and vegetable intakes: impact on attitudes towards dietary change. *British Journal of Nutrition,* 80, 133-140.

Anderson CM, Chambers S, Clamp M, Dunn IA, McGhee MF, Summer KR, and Wood AM (1988). Can audit improve patient care? Effects of studying use of digoxin in general practice. *BMJ,* 297, 113-114.

Anderson GM and Lexchin J (1996). Strategies for improving prescribing practice. *Canadian Medical Association Journal,* 154, 1013-1017.

Angunawela II, Diwan VK, and Tomson G (1991). Experimental evaluation of the effect of drug information on antibiotic prescribing: a study in outpatient care in an area in Sri Lanka. *International Journal of Epidemiology,* 20, 558-564.

Anon. (2002). *Drug formulary and antibiotic policy 2002/2003.* York: York Health Services NHS Trust.

Armitage CJ and Conner M (1999a). Predictive validity of the theory of planned behaviour: the role of questionnaire format and social desirability. *Journal of Community and Applied Social Psychology,* 9, 261-272.

Armitage CJ and Conner M (1999b). The theory of planned behaviour: assessment of predictive validity and 'perceived control. *British Journal of Social Psychology,* 38, 35-54.

Armitage CJ and Conner MT (2001). Efficacy of the theory of planned behaviour: a meta-analytic review. *British Journal of Social Psychology,* 40, 471-499.

Armour BS, Pitts MM, Maclean R, Cangialose C, Kishel M, Imai H, and Etchason J (2001). The effect of explicit financial incentives on physician behavior. *Archives of Internal Medicine,* 161, 1261-1266.

Armstrong D, Reyburn H, and Jones R (1996). A study of general practitioners' reasons for changing their prescribing behaviour. *BMJ,* 312, 949-952.

Asch DA, Jedrziewski M, and Christakis NA (1997). Response rates to mail surveys published in medical journals. *Journal of Clinical Epidemiology,* 50, 1129-1136.

Ashford J, Eccles M, Bond S, Hall L, and Bond J (1999). Improving health care through professional behaviour change: introducing a framework for identifying behaviour change strategies. *British Journal of Clinical Governance,* 4, 14-23.

Ashurst P and Ward D (1983). *An evaluation of counselling in general practice: final report of the Leverhulme Counselling Project.* Mental Health Foundation Report.

Audit Commission (1994). A prescription for improvement: towards more rational prescribing in general practice 1:

Austin SM, Balas EA, and Mitchell J (1994). Effect of physician reminders on preventive care: meta-analysis of randomized clinical trials. *Proceedings the Annual Symposium on Computer Applications in Medical Care,* 121-124.

Aveyard P (1997). Assessing the performance of general practices caring for patients with asthma. *British Journal of General Practice,* 47, 423-426.

Avorn J and Soumerai SB (1983). Improving drug-therapy decisions through educational outreach: a randomized controlled trial of academically based 'detailing'. *New England Journal of Medicine,* 308, 1457-1463.

Axt-Adam P, van der Wouden J, and van der Does E (1993). Influencing behavior of physicians ordering laboratory tests: a literature study. *Medical Care,* 31, 784-794.

Babbie E (1989). *The practice of social research.* 5th Belmon, Calif: Wadsworth.

Bagozzi RP, Baumgartner J, and Yi Y (1989). An investigation into the role of intentions as mediators of the attitude-behavior relationship. *Journal of Economic Psychology,* 10, 35-62.

Baker D and Hann M (2001). General practitioner services in primary care groups in England: is there inequity between service availability and population need? *Health and Place,* 7, 67-74.

Baker R and Fraser RC (1995). Development of review criteria: linking guidelines and assessment of quality. *BMJ,* 311, 370-373.

Baker R, Allen H, Gibson S, Newth J, and Baker E (1996). Evaluation of a primary care counselling service in Dorset. *British Journal of General Practice,* 48, 1049-1053.

Baker R, Lakhani M, Fraser R, and Cheater F (1999). A model for clinical governance in primary care groups. *BMJ,* 318, 779-783.

Balas EA, Boren SA, Brown GD, Ewigman BG, Mitchell JA, and Perkoff GT (1996). Effects of physician profiling on utilization: meta-analysis of randomised clinical trials. *Journal of General Internal Medicine,* 11, 584-590.

Balas EA, Puryear J, Mitchell J, and Barter B (1994). How to structure clinical practice guidelines for continuous quality improvement? *Journal of Medical Systems,* 18: 289-297.

Balas EA, Weingarten S, Grab CT, Blumenthal D, Boren SA, and Brown GD (2000). Improving preventive care by prompting physicians. *Archives of Internal Medicine,* 160, 301-308.

Balk EM, Bonis PAL, Moskowitz H, Schmid CH, Ioannidis JPA, Wang C, and Lau J (2002). Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *JAMA,* 287, 2973-2982.

Banait G, Sibbald B, Thompson D, Summerton C, Hann M, and Talbot S (2003). Modifying dyspepsia management in primary care: a cluster randomised controlled trial of educational outreach compared with passive guideline dissemination. *British Journal of General Practice,* 53, 94-100.

Bandura A (1977a). Self-efficacy: towards a unifying theory of behaviour change. *Psychological Review,* 84, 191-215.

Bandura A (1977b). *Social learning theory.* New Jersey: Prentice Hall.

Bandura A (1998). Health promotion from the perspective of social cognitive theory. *Psychology and Health,* 13, 4-

Bandura A (1986). *Social foundations of thought and action: a social cognitive theory.* New Jersey: Prentice Hall.

Baranowski T, Perry C, and Parcel G (2002). How individuals, environments, and health behavior interact: social cognitive theory. In *Health behavior and health education: theory, research and practice.* 3rd edition, Glanz K, Rimer BK, and Lewis FM (eds). pp 165-184. San Francisco: Jossey Bass.

Barbour RS (2001). Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *BMJ,* 322, 1115-1117.

Barclay S, Todd C, Finlay I, Grande G, and Wyatt P (2002). Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaire studies of GPs. *Family Practice,* 19, 105-111.

Barnard C (1997). The nature of leadership. In *Leadership: classical, contemporary, and critical approaches*. Grint K (eds). pp 89-111. Oxford: Oxford University Press.

Barnes NC, Marone G, Di Maria GU, and et al (1993). A comparison of fluticasone propionate 1 mg daily, with bechlomethasone dipropionate, 2 mg daily in the treatment of asthma. *European Respiratory Journal,* 6, 877-884.

Barnett GO, Winickoff RN, Dorsey JL, Morgan MM, and Lurie SS (1978). quality assurance through automated monitoring and concurrent feedback using a computer-based medical information system. *Medical Care,* 16, 962-970.

Baron RM and Kenny DA (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology,* 51, 1173-1182.

Beaudry JS (1989). The effectiveness of continuing medical education: a quantitative synthesis. *Journal of Continuing Education in the Health Professions,* 9, 285-307.

Becker MH (1974). *The health belief model and personal health behaviour.* Thorofare, New Jersey: Slack.

Beilby JJ and Silagy C (1997). Trials of providing costing information to general practitioners: a systematic review. *Medical Journal of Australia,* 167, 89-92.

Bekker H, Thornton JG, Airey CM, Connelly JB, Hewison J, Robinson MB, Lilleyman J, MacIntosh M, Maule AJ, Michie S, and Pearman AD (1999). Informed decision making: an annotated bibliography and systematic review. *Health Technology Assessment,* 3,

Berings D, Blondeel L, and Habraken H (1994). The effect of industry-independent drug information on the prescribing of benzodiazepines in general practice. *European Journal of Clinical Pharmacology,* 46, 501-505.

Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA, and on behalf of the Cochrane Effective Practice and Organisation of Care Review Group (1998). Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ,* 317, 465-468.

Berwick D (1992). Heal thyself or heal thy system: can doctors help to improve medical care? *Quality in Health Care,* 1 (suppl), 2-8.

Berwick DM (1989). Continuous improvement as an ideal in health care. *New England Journal of Medicine,* 320, 53-56.

Berwick DM (1996). A premier on leading the improvement of systems. *BMJ,* 312, 619-622.

Berwick DM (2003). Disseminating innovations in health care. *JAMA,* 289, 1969-1975.

Bexell A, Lwando E, von Hofsten B, Tembo S, Eriksson B, and Diwan VK (1996). Improving drug use through continuing education: a randomized controlled trial in Zambia. *Journal of Clinical Epidemiology,* 49, 355-357.

Black N (1994). Why we need qualitative research. *Journal of Epidemiology and Community Health,* 48, 425-426.

Black N (2001). Evidence based policy: proceed with care. *BMJ,* 323, 275-279.

Black N and Hutching A (2002). Reduction in the use of surgery for glue ear: did national guideline have an impact? *Quality and Safety in Health Care,* 11, 121-124.

Bland JM and Altman DG (1995). Multiple significance tests: the Bonferroni method. *BMJ,* 310, 170-

Bland JM and Altman DG (1997). Statistics notes: Cronbach's alpha. *BMJ,* 314, 572-

Blank L (2002). Medical professionalism in the new millennium: a physician charter. *Annals of Internal Medicine,* 136, 243-246.

Bloor K and Freemantle N (1996). Lessons from international experience in controlling pharmaceutical expenditure II: influencing doctors. *BMJ,* 312, 1525-1527.

Bloor K, Maynard A, and Freemantle N (1996). Lessons from international experience in controlling pharmaceutical expenditure III: regulating industry. *BMJ,* 313, 33-35.

Blumenthal D and Epstein AM (1996). Quality of health care. Part 6: the role of the physicians in the future of quality management. *New England Journal of Medicine,* 335, 1328-1331.

Boe J, Rosehall L, Alton M, and et al (1989). Comparison of close response effects of inhaled beclamethasone dipropionate and budesonide in the management of asthma. *Allergy,* 44, 349-355.

Bogle SM and Harris CM (1994). Measuring prescribing: the shortcomings of the item. *BMJ,* 308, 637-640.

Bonetti D, Johnston M, Pitts NB, Deery C, Ricketts I, Bahrami M, Ramsay C, and Johnston J (2003). Can psychological models bridge the gap between clinical guidelines and clinicians' behaviour? A randomised controlled trial of an intervention to influence dentists' intention to implement evidence-based practice. *British Dental Journal,* 195, 403-407.

Boot D, Gillies P, Fenelon J, Reubin R, Wilkins M, and Gray P (1994). Evaluation of the short-term impact of counselling in general practice. *Patient Education and Counselling,* 24, 79-89.

Borowitz M and Sheldon T (1993). Controlling health care: from economic incentives to micro-clinical regulation. *Health Economics,* 2, 201-204.

Bottiger LE and Westerholm B (1973). Drug-induced blood dyscrasias in Sweden. *BMJ,* 3, 339-343.

Bouton ME (2000). A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health Psychology,* 19, 57-63.

Bowens A, McDonald R, Ayres P, and Robinson M (2001). The processes and costs of local guideline development: experiences from the Path.Finder NHS consortium. *British Journal of Clinical Governance,* 6, 159-165.

Bower P and Sibbald B (1999). On-site mental health workers in primary care: effects on professional practice. *The Cochrane Database of Systematic Reviews*, Issue 4. Art. No.: CD000532. DOI: 10.1002/14651858.CD000532.

Bower P and Sibbald B (2000). Systematic review of the effect of on-site mental health professionals on the clinical behaviour of general practitioners. *BMJ,* 320, 614-617.

Box GEP and Cox DR (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B: Statistical Methods,* 211-243.

Bradley CP (1992). Turning anecdotes into data - the critical incident technique. *Family Practice,* 9, 98-103.

Brady AJB, Oliver MA, and Pittard JB (2001). Secondary prevention in 24431 patients with coronary heart disease: survey in primary care. *BMJ,* 322, 1463-

Brennan A and Akehurst R (2000). Modelling in health economic evaluation. What is its place? What is its value? *Pharmacoeconomics,* 17, 445-459.

British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, and British Diabetic Association (2000). Joint British recommendations on prevention of coronary heart disease in clinical practice: summary. *BMJ,* 320, 705-708.

British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, and endorsed by the British Diabetic Association (1998). Joint British recommendations on prevention of coronary heart disease in clinical practice. *Heart,* 80, S1-S29.

British Medical Association and Royal Pharmaceutical Society of Great Britain (2001). *British National Formulary: BNF 41.* 41 London: BMJ Books.

British Thoracic Society (1993a). Guidelines for the management of asthma: a summary. *BMJ,* 20, 776-782.

British Thoracic Society (1993b). Guidelines on the management of asthma. Statement by the British Thoracic Society, the Brit. Paediatric Association, the Research Unit of the Royal College of Physicians of London, the King's Fund Centre, the National Asthma Campaign, the Royal College of General Practitioners, the General Practitioners in Asthma Group, the Brit. Assoc. of Accident and Emergency Medicine, and the Brit. Paediatric Respiratory Group. *Thorax,* 48, S1-S24.

British Thoracic Society and National Asthma Campaign (1997a). Asthma in adults and school children. *Thorax,* 52, S1-S21.

British Thoracic Society and National Asthma Campaign (1997b). Asthma in children under five years of age. *Thorax,* 52, S9-S21.

British Thoracic Society and Scottish Intercollegiate Guidelines Network (2003). British Guideline on the Management of Asthma. *Thorax,* 58(Suppl 1),

Britten N (1995). Qualitative interviews in medical research. *BMJ,* 311, 251-253.

Britten N, Stevenson FA, Barry CA, Barber N, and Bradley CP (2000). Misunderstandings in prescribing decisions in general practice: qualitative study. *BMJ,* 320, 484-488.

Britten N and Ukoumunne O (1997). The influence of patients' hopes of receiving a prescription on doctors' perceptions and the decision to prescribe: a questionnaire survey. *BMJ,* 315, 1506-1510.

Brook RH and Williams K (1976). Effect of medical care review on the use of injections: a study of the New Mexico Experimental Care Review Organisation. *Annals of Internal Medicine,* 85, 509-515.

Browner WS, Baron RB, Solkowitz S, Adler LJ, and Gullion DS (1994). Physician management of hypercholesterolemia: a randomized trial of continuing medical education. *Western Journal of Medicine,* 161, 572-578.

Buntinx F, Winkens R, Grol R, and et al (1993). Influencing diagnostic and preventive performance in ambulatory care by feedback and reminders: a review. *Family Practice,* 10, 219-228.

Burgers JS and van Everdingen JJE (2004). Beyond the evidence in clinical guidelines. *Lancet,* 364, 392-393.

Burns TR and Flam H (1987). *The shaping of social organization: Social Rule System Theory with applications*. London: Sage.

Bush P, Rabin DL, and Spector KK (1979). Evaluation of a drug therapy protocol in an HMO. *Medical Care,* 27, 566-577.

Buxton MJ, Drummond MF, Van Hout B, Prince RL, Sheldon TA, Szucs T, and Vray M (1997). Modelling in economic evaluation: an unavoidable fact of life. *Health Economics,* 6, 217-227.

Cabana MD, Rand CS, and et al (2001). Reasons for paediatrician nonadherence to asthma guidelines. *Archives of Paediatrics and Adolescent Medicine,* 155, 1057-1062.

Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PAC, and Rubin HR (1999). Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA,* 282, 1458-1465.

Campbell M, Fitzpatrick R, Haines A, Kinmonth A-L, Sandercock P, Spiegelhalter D, and Tyrer P (2000a). Framework for the design and evaluation of complex interventions to improve health. *BMJ,* 321, 694-696.

Campbell M, Grimshaw J, and Steen N (2000b). Sample size calculations for cluster randomised trials. *Journal of Health Services Research and Policy,* 5, 12-16.

Campbell MJ (2000). Cluster randomized trials in general (family) practice research. *Statistical Methods in Medical Research,* 9, 81-94.

Campbell SM, Cantrill JA, and Roberts D (2000c). Prescribing indicators for UK general practice: Delphi consultation study. *BMJ,* 321, 425-428.

Campbell SM, Hann M, Hacker J, Burns C, Oliver D, Thapar A, Mead N, Gelb Safran D, and Roland MO (2001a). Identifying predictors of high quality care in English general practice: observational study. *BMJ,* 323, 784-787.

Campbell MK, Mollison J, and Grimshaw JM (2001b). Cluster trials in implementation research: estimation of intracluster correlation coefficients and sample size. *Statistics in Medicine*, 20, 391-399.

Cantillon P and Jones R (1999). Does continuing medical education in general practice make a difference? *BMJ,* 318, 1276-1279.

Caro J, Klittich W, McGuire A, Ford I, Norrie J, Pettitt D, McMurray J, and Shepherd J (1997). The West of Scotland coronary prevention study: economic benefit analysis of primary prevention with pravastatin. *BMJ,* 315, 1577-1582.

Carr-Hill R, Jenkins-Clarke S, Dixon P, and Pringle M (1998). Do minutes count? Consultation lengths in general practice. *Journal of Health Services Research and Policy,* 3, 207-213.

Carroll JS and Edmondson AC (2002). Leading organisational learning in health care. *Quality and Safety in Health Care,* 11, 51-56.

Carthy P, Harvey I, Brawn R, and Watkins C (2000). A study of factors associated with cost and variation in prescribing among GPs. *Family Practice,* 17, 36-41.

Centre for Health Services Research (1996). *The primary care management of asthma in adults. North of England evidence based guideline development project.* Newcastle Upon Tyne: Centre for Health Services Research.

Chaftez J (1978). *A premier on the construction of theories in sociology.* Itasca, Ill: Peacock.

Chaix-Couturier C, Durand-Zaleski I, Jolly D, and Durieux P (2000). Effects of financial incentives on medical practice: results from a systematic review of the literature and methodological issues. *International Journal for Quality in Health Care,* 12, 133-142.

Checkland K, Marshall M, and Harrison S (2004). Re-thinking accountability: trust versus confidence in medical practice. *Quality and Safety in Health Care,* 13, 130-135.

Checkland P (1981). *Systems thinking, systems practice.* Chichester: John Wiley & Sons.

Checkland P (1990). *Soft systems methodology in action.* Chichester: John Wiley & Sons.

Choudhry NK, Stelfox HT, and Detsky AS (2002). Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA,* 287, 612-617.

Chrischilles EA and Gondek K (1997). Do practice guidelines augment drug utilisation review? *Pharmacoeconomics,* 12, 648-666.

Clark NM, Gong M, Schork MA, Evans D, Roloff D, Hurwitz M, Maiman L, and Mellins RB (1998). Impact of education for physicians on patient outcomes. *Pediatrics,* 101, 831-836.

Clasen CM, Vernon SW, Mullen PD, and Jackson GL (1994). A survey of physician beliefs and self-reported practices concerning screening for early detection of cancer. *Social Science & Medicine,* 39, 841-849.

Cluzeau F, Littlejohns P, Grimshaw J, Feder G, and Moran SE (1999a). Appraisal instrument for clinical guidelines. *International Journal for Quality in Health Care,* 11, 21-28.

Cluzeau FA, Littlejohns P, Grimshaw JM, Feder G, and Moran SE (1999b). Development and application of a generic methodology to assess the quality of clinical guidelines. *International Journal for Quality in Health Care,* 11, 21-28.

Cockburn J and Pit S (1997). Prescribing behaviour in clinical practice: patients' expectations and doctors' perceptions of patients' expectations. A questionnaire study. *BMJ,* 315, 520-523.

Coe N, Ibbs A, and O'Brien J (1996). *The cost effectiveness of introducing counselling into the primary care setting in Somerset.* Somerset Health Authority Report.

Cohen J (1988). *Statistical power analysis for the behavioral sciences.* Hilhsdale, New Jersey: Lawrence Erlbaum.

Cohen J, Cohen P, West SG, and Aiken L (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* 3rd Mahwah, NJ: Erlbaum.

Comaroff J (1976). A bitter pill to swallow: placebo therapy in general practice. *Sociol Rev,* 24, 79-96.

Community Sport and Recreation Initiatives Project (2000). *Behaviour change theories: social cognitive theory - social learning theory.* Ontario: Parks and Recreation Ontario.

Conner M and Armitage CJ (1998). Extending the theory of planned behavior: a review and avenues for further research. *Journal of Applied Social Psychology,* 28, 1429-1464.

Conner M, Kirk S, Cade JE, and Barrett JH (2001). Why do women use dietary supplements? The use of the theory of planned behaviour to explore beliefs about their use. *Social Science & Medicine,* 52, 621-633.

Conner M and Mcmillan B (1999). Interaction effects in the theory of planned behaviour: studying cannabis use. *British Journal of Social Psychology,* 38, 195-222.

Conner M, Sheeran P, Norman P, and Armitage CJ (2000). Temporal stability as a moderator of relationships in the theory of planned behaviour. *British Journal of Social Psychology,* 39, 469-493.

Conner M and Sparks P (1996). The theory of planned behaviour and health behaviours. In *Predicting health behaviour: research and practice with social cognition models.* Conner M and Norman P (eds). pp 121-162. Buckingham: Open University Press.

Conroy M and Shannon W (1995). Clinical guidelines: their implementation in general practice. *British Journal of General Practice,* 45, 371-375.

Corney R (1984). Social work effectiveness in the management of depressed women: a clinical trial. *Psychological Medicine,* 11, 417-423.

Cox DN, Anderson AS, Reynolds J, McKellar S, Lean M, and Mela D (1998). Take Five, a nutrition education intervention to increase fruit and vegetable intakes: impact on consumer choice and nutrient intakes. *British Journal of Nutrition,* 80, 123-131.

Coyle D, Davies L, and Drummond MF (1998). Trials and tribulations: emerging issues in designing economic evaluations alongside clinical trials. *International Journal of Technology Assessment in Health Care,* 14, 135-144.

Cranney M, Warren E, Barton S, Gardner K, and Walley T (2001). Why do GPs not implement evidence-based guidelines? A descriptive study. *Family Practice,* 18, 359-363.

Crotty M, Whitehead C, Rowett D, Halbert J, Weller D, Finucane P, and Esterman A (2004). An outreach intervention to implement evidence based practice in residential care: a randomized controlled trial [ISRCTN67855475]. *BMC Health Services Research,* 4(6).

Crowne DP and Marlowe D (1960). A new scale of social desirability independent of pschopathology. *Journal of Consulting Psychology,* 24, 349-354.

Curry SJ (2000). Organizational interventions to encourage guideline implementation. *Chest,* 118, 40S-46S.

Davidson R (1998). The transtheoretical model: a critical overview. In *Treating addictive behaviors.* 2nd edition, Miller W and Heather N (eds). pp 25-38. New York: Plenum Press.

Davies HTO and Nutley S (2000). Developing learning organisations in the new NHS. *BMJ,* 320, 998-1001.

Davis D, Thomson O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, and Taylor-Vaisey A (1999). Impact of formal continuing medical education. Do conferences, workshops, rounds, and other traditional continuing education activities change physician behaviour or health outcome? *JAMA,* 282, 867-874.

Davis DA and Taylor-Vaisey A (1997). Translating guidelines into practice: a systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. *Canadian Medical Association Journal,* 157, 408-416.

Davis DA, Thomson MA, Oxman AD, and Haynes B (1992). Evidence for the effectiveness of CME: a review of 50 randomized controlled trials. *JAMA,* 268, 1111-1117.

Davis DA, Thomson MA, Oxman AD, and Haynes B (1995). Changing physician performance: a systematic review of the effect of continuing medical education strategies. *JAMA,* 274, 700-705.

Davis D, Evans M, Jadad A, Perrier L, Rath D, Ryan D, Sibbald G, Straus S, Rappolt S, Wowk M, and Zwarenstein M (2003). The case for knowledge translation: shortening the journey from evidence to effect. *BMJ,* 327, 33-35.

de Burgh S, Mant A, Mattick RP, Donnelly N, Hall W, and Bridges-Webb C (1995). A controlled trial of educational visiting to improve benzodiazepine prescribing in general practice. *Australian Journal of Public Health,* 19, 142-148.

De Santis G, Harvey KJ, Howard D, Mashford ML, and Moulds R (1994). Improving the quality of antibiotic prescription patterns in general practice: the role of educational intervention. *Medical Journal of Australia,* 160, 502-505.

Denig P, Haaijer-Ruskamp FM, and Zijsling DH (1990). Impact of a drug bulletin on the knowledge, perception of drug utility, and prescribing behavior of physicians. *DICP, Annals of Pharmacotherapy,* 24, 87-93.

Denig P, Witteman CLM, and Schouten HW (2002). Scope and nature of prescribing decisions made by general practitioners. *Quality and Safety in Health Care,* 11, 137-143.

Department of Health (1998). *A first class service: quality in the new NHS.* London: Department of Health.

Department of Health (2001a). *National Service Framework for older people.* London: Department of Health.

Department of Health (2001b). *NHS performance indicators: a consultation.* London: Department of Health.

Department of Health (2001c). *Shifting the balance of power within the NHS: securing delivery.* London: Department of Health.

Di Iorio D, Henley E, and Doughty A (2000). A survey of primary care physician practice patterns and adherence to acute low back problem guidelines. *Archives of Family Medicine,* 9, 1015-1021.

Diamond GA (1989). Through a lens, darkly. *?,*

Dickinson J, Hutton S, and Atkin A (1998). Implementing the British Thoracic Society's guidelines: the effect of a nurse-run asthma clinic on prescribed treatment in an English general practice. *Respir Med,* 92, 264-267.

Dickinson JC, Warshaw GA, Gehlbach SH, Bobula JA, Muhlbaier LH, and Parkerson GR (1981). Improving hypertension control: impact of computer feedback and physician education. *Medical Care,* 19, 843-854.

DiClemente CC and Prochaska JO (1998). Towards a comprehensive, transtheoretical model of change: stages of change and addictive behaviour. In *Treating addictive*

behaviors. 2nd edition, Miller W and Heather N (eds). pp 3-24. New York: Plenum Press.

Dinarvand R and Nikzad A (2000). Status of prescription and drug usage in Tehran in 1998. *Hakim,* 3,

Diwan VK, Wahlstrom R, Tomson G, Beermann B, Sterky G, and Eriksson B (1995). Effects of "group detailing" on the prescribing of lipid-lowering drugs: a randomized controlled trial in Swedish primary care. *Journal of Clinical Epidemiology,* 48, 705-711.

Dobbins M, Cockerill R, Barnsley J, and Ciliska D (2001). Factors of the innovation, organization, environment, and individual that predict the influence five systematic reviews had on public health decisions. *International Journal of Technology Assessment in Health Care,* 17, 467-478.

Dolan P and Torgerson D (1998). The cost of treating osteoporosis fractures in the UK female population. *Osteoporosis International,* 8, 611-617.

Donabedian A (2003). *An introduction to quality assurance in health care.* Oxford: Oxford University Press.

Donaldson LJ and Donaldson RJ (2003). The National Health Service. In *Essential Public Health.* 2nd edition, pp 169-235. Plymouth: Petroc Press.

Dowding D (2001). Examining the effects that manipulating information given in the change of shift report has on nurses' care planning ability. *Journal of Advanced Nursing,* 33, 836-846.

Dowding D and Thompson C (2003). Measuring the quality of judgement and decision-making in nursing. *Journal of Advanced Nursing,* 44, 49-57.

Dowie R (1998). A review of research in the United Kingdom to evaluate the implementation of clinical guidelines in general practice. *Family Practice,* 15, 462-470.

Dowswell G, Harrison S, and Wright J (2001). Clinical guidelines: attitudes, information processes and culture in English primary care. *International Journal of Health Planning and Management,* 16, 107-124.

Drummond MF, O'Brien BJ, Stoddert GL, and Torrance GW (1997). *Methods for the economic evaluation of health care programmes.* 2 Oxford: Oxford University Press.

Duran-Tauleria E and Rona R (1999). Geographical and socioeconomic variation in the prevalence of asthma symptoms in English and Scottish children. *Thorax,* 54, 476-481.

Durieux P, Ravaud P, Chaix C, and Durand-Zaleski I (1999). Does continuing medical education improve the way physicians conduct their practice? *Presse Medicale,* 28, 468-472.

Durieux P, Ravaud P, Dosquet P, and Durocher A (2000). Mise en oeuvre des recommandations medicales : revue systematique des revues systematiques. *Gastroenterologie Clinique et Biologique,* 24, 1018-1025.

Earll L and Kincey J (1982). Clinical psychology in general practice: a controlled trial evaluation. *Journal of the Royal College of General Practitioners,* 32, 32-37.

Eccles M, Grimshaw J, Johnston M, Pitts NB, and Walker A (2001a). *Using health psychology models in implementation reseach: planned investigations (PRIME and INTERIM) [abstract].* St Andrews, Scotland: Proceeding of the British Psychological Society: Proceedings from the European Health Psychology Society Annual Conference, Changing Behaviour: Health and Healthcare.

Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, and Purves I (2002). Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ,* 325, 941-944.

Eccles M, Rousseau N, Higgins B, Thomas L, and for the North of England Asthma Guideline Development Group (2001b). Evidence-based guideline on the primary care management of asthma. *Family Practice,* 18, 223-229.

Economic & Social Research Council Research Methods Programme (2003). http://www.ccsr.ac.uk/methods.

Eddy DM (1982). Clinical policies and the quality of clinical practice. *New England Journal of Medicine,* 307, 343-347.

Eddy DM (1984). Variations in physician practice: the role of uncertainty. *Health Affairs,* 3, 74-89.

Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, and Kwan I (2002). Increasing response rates to postal questionnaires: systematic review. *BMJ,* 324, 1183-1185.

Eldridge S, Cryer C, Feder G, and Underwood M (2001). Sample size calculations for intervention trials in primary care randomizing by primary care group: an empirical illustration from one proposed intervention trial. *Statistics in Medicine,* 20, 367-376.

Ellison MD, Edwards LB, Edwards EB, Barker CF (2003). Geographic differences in access to transplantation in the United States. *Transplantation,* 76, 1389-1394.

Engestrom Y (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics,* 43, 960-974.

Escarce JJ (1996). Externalities in hospitals and physician's adoption of a new surgical technology: an exploratory analysis. *Journal of Health Economics,* 15, 715-734.

Evans CE, Haynes RB, Birkett NJ, Gilbert JR, Taylor DW, Sackett DL, Johnston ME, and Hewston SA (1986). Does a mailed continuing education program improve physician performance? Results of a randomized trial in antihypertensive care. *JAMA,* 255, 501-504.

Ezzo J, Bausell B, Moerman DE, Berman B and Hadhazy V (2001). Reviewing the reviews: how strong is the evidence? How clear are the conclusions? *International Journal of Technology Assessment in Health Care,* 17, 457-466.

Fairhurst K and Huby G (1998). From trial data to practical knowledge: qualitative study of how general practitioners have accessed and used evidence about statin drugs in their management of hypercholestrolemia. *BMJ,* 317, 1130-1134.

Feder G, Eccles M, Grol R, Griffiths C, and Grimshaw J (1999). Clinical guidelines: using clinical guidelines. *BMJ,* 318, 728-730.

Feder G, Griffiths C, Highton C, Eldridge S, Spence M, and Southgate L (1995). Do clinical guidelines introduced with practice based education improve care of asthmatic and

diabetic patients? A randomised controlled trial in general practice in east London. *BMJ,* 311, 1473-1478.

Ferlie E (1997). Large scale organisational and managerial change in health care: a review of the literature. *Journal of Health Services Research and Policy,* 2, 180-188.

Ferlie E (2001). Organisational studies. In *Studying the organisation and delivery of health services: research methods.* Fulop N, Allen P, Clarke A, and Black N (eds). pp 24-40. London: Routledge.

Ferlie E, Gabbay J, Fitzgerald L, Locock L, and Dopson S (2001). Evidence-based medicine and organisational change: an overview of some recent qualitative research. In *Organisational behaviour and organisational studies in health care: reflections on the future.* Ashburner L (eds). pp 18-42. New York: Palgrave.

Field MJ and Lohr KN (1990). *Clinical practice guidelines: directions for a new program.* Washington: National Academies Press.

Flottorp S, Oxman A, and Bjorndal A (1998). The limits of leadership: opinion leaders in general practice. *Journal of Health Services Research and Policy,* 3, 197-202.

Folland S, Goodman AC, and Stano M (2004). *The economics of health and health care.* 4th New Jersey: Pearson Prentice Hall.

Foreman DM, Foreman D, and Minty E (2003). How should we measure social disadvantage in clinic setting? *European Child and Adolescent Psychiatry,* 12, 308-312.

Freeman AC and Sweeney K (2001). Why general practitioners do not implement evidence: qualitative study. *BMJ,* 323, 1100-1102.

Freemantle N (1996). Are decisions taken by health care professionals rational? A non systematic review of experimental and quasi experimental literature. *Health Policy,* 38, 71-81.

Freemantle N and Bloor K (1996). Lessons from international experience in controlling pharmaceutical expenditure. I: influencing patients. *BMJ,* 312, 1469-1471.

Freemantle N, Harvey EL, Wolf F, Grimshaw JM, Grilli R, and Bero LA (1997). Printed educational materials: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD000172. DOI: 10.1002/14651858.CD000172.

Freemantle N, Eccles M, Mason J, Thomson MA, Wolf FM, and Wood J (1998). Research implementation methods. In *Health services research: a guide to best practice.* 1 edition, Black N, Brazier J, Fitzpatrick R, and Reeves B (eds). pp 237-246. London: BMJ Books.

Freemantle N, Johnson R, Dennis J, Kennedy A, and Marchment M (2000). Sleeping with the enemy? A randomized controlled trial of a collaborative health authority/industry intervention to influence prescribing practice. *British Journal of Clinincal Pharmacology,* 49, 174-179.

Freemantle N, Nazareth I, Eccles M, Wood J, and Haines A (2002). A randomised trial of the effect of educational outreach by community pharmacists on prescribing in UK general practice. *British Journal of General Practice,* 52, 290-295.

French DP and Hankins M (2003). The expectancy–value muddle in the theory of planned behaviour — and some proposed solutions. *British Journal of Health Psychology,* 8, 37-55.

Freudenstein U and Howe A (1999). Recommendations for future studies: a systematic review of educational interventions in primary care settings. *British Journal of General Practice,* 49, 995-1001.

Frischer M, Heatlie H, Chapman S, Norwood J, Bashford J, and Millson D (1999). Should the corticosteroid to bronchodilator ratio be promoted as a quality prescribing marker? *Public Health,* 113, 247-250.

Fuchs VR (2000). The future of health economics. *Journal of Health Economics,* 19, 141-157.

Garner P, Kale R, Dickson R, Dans T, and Salinas R (1998). Getting research findings into practice: implementing research findings in developing countries. *BMJ,* 317, 531-535.

Gehlbach SH, Wilkinson WE, Hammond WE, Clapp NE, Finn AL, Taylor WJ, and Rodell MS (1984). Improving drug prescribing in a primary care practice. *Medical Care,* 22, 193-201.

Giacomini MK, Cook DJ, for the Evidence-Based Medicine Working Group (2000). Users' Guides to the medical literature, XXIII: qualitative research in health care. A. are the results of the study valid? *JAMA*, 284, 357-362.

Gill PS, Makela M, Vermeulen KM, Freemantle N, Ryan G, Bond C, Thorsen T, and Haaijer-Ruskamp FM (1999). Changing doctor prescribing behaviour. *Pharmacy World and Science,* 21, 158-167.

Ginsberg G, Marks I, and Waters H (1984). Cost-benefit analysis of a controlled trial of nurse therapy for neuroses in primary care. *Psychological Medicine,* 14, 683-690.

Giuffrida A, Gosden T, Forland F, Kristiansen I, Sergison M, Leese B, Pedersen L, and Sutton M (1999). Target payments in primary care: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews.* Issue 4. Art. No.: CD000531. DOI: 10.1002/14651858.CD000531.

Glanz K and Maddock J (2000). On judging the models and theories: research and practice, psychology and public health. *Journal of Health Psychology,* 5, 151-154.

Glanz K, Rimer BK, and Lewis FM (2002a). *Health behavior and health education: theory, research and practice.* 3rd San Francisco: Jossey Bass.

Glanz K, Rimer BK, and Lewis FM (2002b). Theory, research, and practice in health behaviour and health education. In *Health behavior and health education: theory, research and practice.* 3rd edition, Glanz K, Rimer BK, and Lewis FM (eds). pp 22-39. San Francisco: Jossey Bass.

Godin G and Kok G (1996). The theory of planned behavior: a review of its application to health behaviors. *American Journal of Health Promotion,* 11, 87-98.

Goldberg HI, Wagner EH, Fihn SD, Martin DP, Horowitz CR, Christensen DB, Cheadle AD, Diehr P, and Simon G (1998). A randomized controlled trial of CQI teams and academic detailing: can they alter compliance with guidelines? *Joint Commission Journal on Quality Improvement,* 24, 130-142.

Gollwitzer PM (1999). Implementation intentions: strong effects of simple plans. *American Psychologist,* 54, 493-503.

Gollwitzer PM and Oettingen G (2000). The emergence and implementation of health goals. In *Understanding and changing health behaviour: from health beliefs to self-*

regulation. Norman P, Abraham C, and Conner M (eds). pp 229-260. Amsterdam: Harwood Academic Publishers.

Gonzales R, Steiner JF, Lum A, and Barrett PHJr (1999). Decreasing antibiotic use in ambulatory practice: impact of a multidimentional intervention on the treatment of uncomplicated acute bronchitis in adults. *JAMA,* 281, 1512-1519.

Gorton TA, Cranford CO, Golden WE, Walls RC, and Pawelak JE (1995). Primary care physicians' response to dissemination of practice guidelines. *Archives of Family Medicine,* 4, 135-142.

Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Giuffrida A, Sergison M, and Pedersen L (2000). Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. *The Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD002215. DOI: 10.1002/14651858.CD002215.

Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Giuffrida A, Sergison M, and Pedersen L (2001). Impact of payment method on behaviour of primary care physicians: a systematic review. *Journal of Health Services Research and Policy,* 6, 44-55.

Gosden T, Pedersen L, and Torgeson D (1999). How should we pay doctors? A systematic review of salary payments and their effect on doctor behaviour. *Quarterly Journal of Medicine,* 92, 47-55.

Grad R, Tamblyn R, McLeod P, Snell L, Illescas A, and Boudreau D (1997). Does knowledge of drug prescribing predict drug management of standardized patients in office practice? *Medical Education,* 31, 132-137.

GRADE Working Group (2004). Grading quality of evidence and strength of recommendations. *BMJ,* 328, 1490-1494.

Graham ID, Beardall S, Carter AO, and et al (2001). What is the quality of drug therapy clinical practice guidelines in Canada? *Canadian Medical Association Journal,* 165, 157-163.

Graham ID, Calder LA, Hebert PC, Carter AO, and Tetroe JM (2000a). A comparison of clinical practice guideline appraisal instruments. *International Journal of Technology Assessment in Health Care,* 16, 1024-1038.

Graham RP, James P, and Cowan TM (2000b). Are clinical practice guidelines valid for primary care? *Journal of Clinical Epidemiology,* 53, 949-954.

Greco PJ and Eisenberg JM (1993). Changing physicians' practice. *New England Journal of Medicine,* 329, 1271-1273.

Green J (2000). Epistemology, evidence and experience: evidence based health care in the work of Accident Alliances. *Sociology of Health & Illness,* 22, 453-476.

Green L, Kreuter M, Deeds S, and Partridge K (1980). *Health education planning: a daignostic approach.* Palo Alto, California: Mayfield Press.

Green SB (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research,* 26, 499-510.

Greenhalgh J and Meadows K (1999). The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *Journal of Evaluation in Clinical Practice,* 5, 401-416.

Greenhalgh T and Taylor R (1997). How to read a paper: papers that go beyond numbers (qualitative research). *BMJ*, 315, 740-743.

Greer AL (1977). Advances in the study of diffusion of innovation in health care organizations. *Milbank Memorial Fund Quarterly,* 55, 505-532.

Greer AL (1985). Adoption of medical technology. The hospital's three decision systems. *International Journal of Technology Assessment in Health Care,* 1, 669-680.

Greer AL (1988). The state of art versus the state of the science: the diffusion of new medical technologies into practice. *International Journal of Technology Assessment in Health Care,* 4, 5-26.

Greer AL (1995). The shape of resistance, the shapers of change. *Joint Commission Journal on Quality Improvement,* 21, 328-332.

Greer AL, Goodwin JS, Freeman JL, and Wu ZH (2002). Bringing the patient back in: guidelines, practice variations, and the social context of medical practice. *International Journal of Technology Assessment in Health Care,* 18, 747-761.

Griffiths C, Kaur G, Gantley M, Feder G, Hillier S, Goddard J, and Packe G (2001). Influences on hospital admission for asthma in south Asian and white adults: qualitative interview study. *BMJ,* 323, 962-966.

Griffiths C, Naish J, Sturdy P, and Pereira F (1996). Prescribing and hospital admissions for asthma in east London. *BMJ,* 312, 481-482.

Griffiths C, Sturdy P, Naish J, Omar R, Dolan S, and Feder G (1997). Hospital admissions for asthma in East London: association with characteristics of local general practices, prescribing, and population. *BMJ,* 314, 482-486.

Grilli R and Lomas J (1994). Evaluating the message: the relationship between compliance rate and the subject of a practice guideline. *Medical Care,* 32, 202-213.

Grilli R, Magrini N, Penna A, Mura G, and Liberati A (2000). Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet,* 355, 103-106.

Grilli R, Ramsay C, and Minozzi S (2002). Mass media interventions: effects on health services utilisation. *The Cochrane Database of Systematic Reviews*, Issue 1. Art. No.: CD000389. DOI: 10.1002/14651858.CD000389.

Grimes DA and Schulz KF (2002). Bias and causal associations in observational research. *Lancet,* 359, 248-252.

Grimshaw JM (1998). *Evaluation of four quality assurance initiatives to improve out-patient referrals from general practice to hospital.* PhD Thesis. Aberdeen: University of Aberdeen, 1998.

Grimshaw J, Freemantle N, Wallace S, Russell I, Hurwitz B, Watt I, Long A, and Sheldon T (1995). Developing and implementing clinical practice guidelines. *Quality in Health Care,* 4, 55-64.

Grimshaw J, McAuley LM, Bero LA, Grilli R, Oxman AD, Ramsay C, Vale L, and Zwarenstein M (2003). Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Quality and Safety in Health Care*, 12, 298-303.

Grimshaw JM and Russell IT (1993). Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet,* 342, 1317-1322.

Grimshaw JM and Russell IT (1994). Achieving health gain through clinical guidelines II: ensuring guidelines change medical practice. *Quality in Health Care,* 3, 45-52.

Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, Grilli R, Harvey E, Oxman A, and O'Brien MA (2001). Changing provider behaviour: an overview of systematic reviews of interventions. *Medical Care,* 39, II2-II45.

Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale R, Whitty P, Eccles MP, Matowe L, Shirran L, Wensing M, Dijkstra R, and Donaldson C (2004). Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technology Assessment,* 8(6), 1-84.

Grol R (1992). Implementing guidelines in general practice care. *Quality in Health Care,* 1, 184-191.

Grol R (1997). Beliefs and evidence in changing practice. *BMJ,* 315, 418-421.

Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, and Mokkink H (1998). Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *BMJ,* 317, 858-861.

Grol R and Jones R (2000). Twenty years of implementation research. *Family Practice,* 17, S32-S35.

Grol R (2001). Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Medical Care,* 39, II46-II54.

Gross PA and Pujat D (2001). Implementing practice guidelines for appropriate antimicrobial usage: a systematic review. *Medical Care,* 39, II55-II69.

Gurwitz JH, Soumerai SB, and Avorn J (1990). Improving medication prescribing and utilisation in the nursing home. *J Am Geriatr Soc,* 38, 542-552.

Gustafsson A, Herrmann A, and Huber F (2003). *Conjoint measurement: methods and applications.* 3rd Berlin: Springer.

Gutiérrez G, Guiscafré H, Bronfman M, Walsh J, Martinez H, and Munoz O (1994). Changing physician prescribing patterns: evaluation of an educational strategy for acute diarrhea in Mexico City. *Medical Care,* 32, 436-446.

Gyorkos TW, Tannenbaum T, Abrahamowicz M, and et al (1994). Evaluation of the effectiveness of immunization delivery methods. *Canadian Journal of Public Health,* 85, S14-S30.

Hadiyono JEP, Suryawati S, Danu SS, Surantono, and Santoso B (1996). Interactional group discussion: results of a controlled trial using a behavioral intervention to reduce the use of injections in public health facilities. *Social Science & Medicine,* 42, 1177-1183.

Hafeez A, Kiani AG, ud Din S, Muhammad W, Butt K, Shah Z, and Mirza Z (2004). Prescription and dispensing practices in public sector health facilities in Pakistan: survey report. *Journal of Pakistani Medical Association,* 54, 187-191.

Hag IU, Jackson PR, Yeo WW, and Ramsay LE (1995). Sheffield risk and treatment table for cholestrol lowering for primary prevention of coronary heart disease. *Lancet,* 346, 1467-1471.

Haines A and Jones R (1994). Implementing findings of research. *BMJ,* 308, 1488-1492.

Ham C (2003). Improving the performance of health services: the role of clinical leadership. *Lancet,* 361, 1978-1980.

Hammond KR, Rohrbaugh J, Mumpower J, and Adelman L (1977). Social judgement theory: applications in policy formation. In *Human judgement and decision processes in applied settings.* Kaplan MF and Schwartz S (eds). pp 1-27. New York: Academic Press.

Hankins M, French D, and Horne R (2000). Statistical guidelines for studies of the theory of reasoned action and the theory of planned behaviour. *Psychology and Health,* 15, 151-161.

Hansell A, Hollowell J, Nicholas T, McNiece R, and Strachan D (1999). Use of the General Practice Research Database (GPRD) for respiratory epidemiology: a comparison with the 4th Morbidity Survey in Genral Practice (MSGP4). *Thorax*, 54, 413-419.

Hanson LC, Tulsky JA, and Danis M (1997). Can clinical interventions change care at the end of life? *Annals of Internal Medicine,* 126, 381-388.

Hardeman W, Johnston M, Johnston D, Bonetti D, Wareham N, and Kinmonth A (2002). Application of theory of planned behavior in behavior change interventions: a systematic review. *Psychology and Health,* 17, 123-158.

Harris C, Fry J, Jarman B, and Woodman E (1985). Prescribing - a case for prolonged treatment. *Journal of the Royal College of General Practitioners,* 35, 284-287.

Harris CM, Jarman B, Woodman E, White P, and Fry JS (1984). Prescribing - a suitable case for treatment. *Journal of the Royal College of General Practitioners - Occasional Paper,* 24, 1-39.

Harrison S, Dowswell G, and Wright J (2002). Practice nurses and clinical guidelines in a changing primary context: an empirical study. *Health and Nursing Policy Issues*, 39, 299-307.

Harrison S, Dowswell G, Wright J, and Russell I (2003). General practitioners' uptake of clinical guidelines: a qualitative study. *Journal of Health Services Research and Policy,* 8, 149-153.

Hart O (1995). An economist's perspective on the theory of the firm. In *Organizationn theory: from Chester Barnard to the present and beyond.* 2nd edition, Williamson OE (eds). pp 154-171. New York: Oxford University Press.

Hasenfeld R and Shekelle PG (2003). Is the methodological quality of guidelines declining in the US? Comparison of the quality of US Agency for Health Care Policy and Research (AHCPR) guidelines with those published subsequently. *Quality and Safety in Health Care,* 12, 428-434.

Hayduk L (1987). *Structural equation modeling with LISREL.* Baltimore, Maryland: John Hopkins.

Haynes RB, Davis DA, McKibbon A, and Tugwell P (1984). A critical appraisal of the efficacy of continuing medical education. *JAMA,* 251, 61-64.

Haynes RB and Walker CJ (1987). Computer-aided quality assurance: a critical appraisal. *Archives of Internal Medicine,* 147, 1297-1301.

Hemmings A (1997). Counselling in primary care: a randomised controlled trial. *Patient Education and Counselling,* 32, 219-230.

Hetlevik I, Holmen J, Krüger O, Kristensen P, Iversen H, and Furuseth K (2000). Implementing clinical guidelines in the treatment of diabetes mellitus in general practice: evaluation of effort, process, and patient outcome related to implementation of a computer-based decision support system. *International Journal of Technology Assessment in Health Care,* 16, 210-227.

Hibble A, Kanka D, Pencheon D, and Pooles F (1998). Guidelines in general practice: the new Tower of Babel? *BMJ,* 317, 862-863.

Hicks L (2002). The nature of learning. In *Management and organisational behaviour.* 6th edition, Mullins L (eds). pp 349-384. Harlow, England: Prentice Hall.

Hippisley-Cox J, Pringle M, Coupland C, Hammersley V, and Wilson A (2001). Do single handed practices offer poorer care? Cross sectional survey of processes and outcomes. *BMJ,* 323, 320-323.

Hiss RG, MacDonald R, and David WR (1978). *Identification of physician educational influentials in small community hospitals.* Washington DC: Proceedings of the Seventeenth Annual Conference on Research in Medical Education.

Hobbs FD, Delaney BC, Carson A, and Kenkre JE (1996). A prospective controlled trial of computerized decision support for lipid management in primary care. *Family Practice,* 13, 133-137.

Hogerzeil HV, Walker GJ, Sallami AO, and Fernando G (1989). Impact of an essential drug programme on availability and rational use of drugs. *Lancet,* January 21, 141-142.

Holloway K (2004). Rational use of drugs: an overview

Hopayian K (2001). The need for caution in interpreting high quality systematic reviews. *BMJ,* 323, 681-684.

Horder J, Bosanquet N, and Stocking B (1986). Ways of influencing the behaviour of general practitioners. *Journal of Royal College of General Practitioners,* 36, 517-521.

Howard P and Duncan PW (1997). Primary stroke prevention in nonvalvular atrial fibrillation: implementing the clinical trial findings. *Annals of Pharmacotherapy,* 31, 1187-1196.

Howitt A and Armstrong D (1999). Implementing evidence based medicine in general practice: audit and qualitative study of antithrombotic treatment for atrial fibrillation. *BMJ,* 318, 1324-1327.

Hsieh F, Bloch DA, and Larsen M (1998). A simple method of sample size calculation for linear and logistic regression. *Statistics in Medicine,* 17, 1623-1634.

Hulscher MEJL, Wensing M, van der Weijden T, and Grol R (1997). Interventions to implement prevention in primary care. *The Cochrane Database of Systematic Reviews*, Issue . Art. No.: CD000362. DOI: 10.1002/14651858.CD000362.

Hulscher M (1998). Implementing prevention in general practice: a study on cardiovascular disease [PhD thesis]

Hunt DL, Haynes RB, Hanna SE, and Smith K (1998). Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *JAMA,* 280, 1339-1346.

Hux JE, Melady MP, and DeBoer D (1999). Confidential prescriber feedback and education to improve antibiotic use in primary care: a controlled trial. *Canadian Medical Association Journal,* 161, 388-392.

Iglehart JK (1991). Efforts to address the problem of physician self-referral. *New England Journal of Medicine,* 325, 1820-1823.

Iles V and Sutherland K (2001). *Managing change in the NHS. Organisational change: a review for health care managers, professionals and researchers.* London: National Co-ordinating Centre for Service Delivery and Organisation.

Ilett KF, Johnson S, Greenhill G, Mullen L, Brockis J, Golledge CL, and Reid DB (2000). Modification of general practitioner prescribing of antibiotics by use of a therapeutics adviser (academic detailer). *British Journal of Clinincal Pharmacology,* 49, 168-173.

Inman WHW and Adelstein AM (1969). Rise and fall of asthma mortality in England and Wales in relation to use of pressurised aerosols. *Lancet,* 2, 279-285.

Irani J, Brown CT, van der Meulen J, and Emberton M (2003). A review of guidelines on benign prostatic hyperplasia and lower urinary tract symptoms: are all guidelines the same? *BJU International,* 92, 937-942.

Jack W (2000). *Principles of health economics for developing countries.* 1 Washington DC: World Bank.

Jackson R (2000). Updated New Zealand cardiovascular disease risk-benefit prediction guide. *BMJ,* 320, 709-710.

Jacoby A, Smith M, and Eccles M (2003). A qualitative study to explore influences on general practitioners' decisions to prescribe new drugs. *British Journal of General Practice,* 53, 120-125.

Jadad AR, Cook DJ, and Browman GP (1997). A guide to interpreting discordant systematic reviews. *Canadian Medical Association Journal,* 156, 1411-1416.

Jamtvedt G, Young JM, Kristoffersen DT, Thomson O'Brien MA, and Oxman AD (2003). Audit and feedback: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD000259. DOI: 10.1002/14651858.CD000259.

Jarman B (1983). Identification of underprivileged areas. *BMJ,* 286, 1705-1709.

Jarman B (1984). Underprivileged areas: validation and distribution of scores. *BMJ,* 289, 1587-1592.

Jaye C (2002). Doing qualitative research in general practice: methodological utility and engagement. *Family Practice*, 19, 557-562.

Jefferys M and Sachs H (1983). *Rethinking general practice: dilemmas in primary medical care.* London: Tavistock.

Jemmott J, Jemmott L, and Fong GT (1998). Abstinence and safer sex HIV risk-reduction interventions for African American adolescents: a randomized controlled trial. *JAMA*, 279, 1529-1536.

Jensen UJ and Mooney G (1990). Changing values: autonomy and paternalism in medicine and health care. In *Changing values in medical and health care decision making.* Jensen UJ and Mooney G (eds). pp 1-15. Chichester, England: John Wiley & Sons.

Johnson RA, Kast FE, and Rosenzweig JE (1973). *The theory and management of systems.* 3rd New York: McGraw-Hill.

Johnston G, Crombie IK, Davies HTO, Alder EM, and Millard A (2000). Reviewing audit: barriers and facilitating factors for effective clinical audit. *Quality in Health Care,* 9, 23-36.

Johnston ME, Langton KB, Haynes RB, and Mathieu A (1994). Effects of computer-based clinical decision support systems on clinician performance and patient outcome: a critical appraisal of research. *Annals of Internal Medicine,* 120, 135-142.

Jonas K (1993). Expectancy-value models of health behaviour: an analysis by conjoint measurement. *European Journal of Social Psychology,* 23, 167-183.

Jones DL, Kroenke K, Landry FJ, Tomich DJ, and Ferrel RJ (1996). Cost savings using a stepped-care prescribing protocol for nonsteroidal anti-inflammatory drugs. *JAMA,* 275, 926-930.

Jones JW (2000). The healthcare professional and the Bolam test. *British Dental Journal,* 188, 237-240.

Jones K (1992). Impact of an interest in asthma on prescribing costs in general practice. *Quality in Health Care,* 1, 110-113.

Jones KP, Harris CM, Bogle SM, Foley R, and Usherwood TP (1995). The effects on prescribing patterns and costs of having a special interest in asthma. *Journal of Royal Society of Medicine,* 88, 570-575.

Jones MI, Greenfield SM, Jowett S, and et al (2001). Proton pump inhibitors: a study of GPs' prescribing. *Family Practice,* 18, 333-338.

Jones P, Kafonek S, Laurora I, Hunninghake D, and for the CURVES Investigators (1998). Comparative dose efficacy study of Atorvastatin versus Simvastatin, Pravastatin, Lovastatin, and Fluvastatin in patients with hypercholesterolemia (the CURVES Study). *American Journal of Cardiology,* 81, 582-587.

Jones RH, Lydread S, and Dunleavey J (1993). Problems with implementing guidelines: a randomised controlled trial of consensus management of dyspepsia. *Quality in Health Care,* 2, 217-221.

Juni P, Witschi A, Bloch R, and Egger M (1999). The hazards of scoring the quality of clinical trial for meta-analysis. *JAMA,* 282, 1054-1060.

Kafuko JM, Zirabamuzaale, and Bagena D (1999). *Rational drug use in rural health units of Uganda:effect of national standard treatment guidelines on rational drug use.* 1st International Conference on Improving Use og Medications.

Kaner EF, Haighton CA, and McAvoy BR (1998). 'So much post, so busy with practice--so, no time!': a telephone survey of general practitioners' reasons for not participating in postal questionnaire surveys. *British Journal of General Practice,* 48, 1067-1069.

Kanouse DE and Jacoby I (1988). When does information change practitioners' behavior? *International Journal of Technology Assessment in Health Care,* 4, 27-33.

Katon W, Robinson P, von Korff M, Lin E, Bush T, Ludman E, and et al (1996). A multifaceted intervention to improve treatment of depression in primary care. *Archives of General Psychiatry,* 53, 924-932.

Katon W, Von Korff M, Lin E, Walker E, Simon GE, Bush T, and et al (1995). Collaborative management to achieve treatment guidelines: impact on depression in primary care. *JAMA,* 273, 1026-1031.

Katon W, von Korff M, Lin E, Bush T, and Ormel J (1992a). Adequacy and duration of antidepressant treatment in primary care. *Medical Care,* 30, 67-76.

Katon W, von Korff M, Lin E, Bush T, Russo J, Lipscomb P, and et al (1992b). A randomized trial of psychiatric consultation with distressed high utilizers. *General Hospital Psychiatry,* 14, 86-98.

Kaufman DM (2003). ABC of learning and teaching in medicine: applying educational theory in practice. *BMJ,* 326, 213-216.

Kerlinger F (1986). *Foundations of behavioral research.* 3rd New York: Holt, Rinehart and Winston.

Kerry SM and Bland JM (1998). The intracluster coefficient in cluster randomisation. *BMJ,* 316, 1455-1460.

Kerwick SW and Jones RH (1996). Educational interventions in primary care psychiatry: a review. *Primary Care Psychiatry,* 2, 107-117.

King DJ, Griffiths K, Hall CE, Cooper NC, and Morrison DJ (1980). Effects of the CURB campaign on barbiturate prescribing in Northern Ireland. *Journal of the Royal College of General Practitioners,* 30, 614-618.

Kitson A, Harvey G, and McKormack B (1998). Enabling the implementation of evidence based practice: a conceptual framework. *Quality in Health Care,* 7, 149-158.

Klein LE, Charache P, and Johannes RS (1981). Effect of physician tutorials on prescribing patterns of graduate physicians. *Journal of Medical Education,* 56, 504-511.

Kotler P and Roberto E (1989). *Social marketing: strategies for changing public behavior.* New York: Free Press.

Krasnik A, Groenewegen PP, Pedersen PA, von Scholten P, Mooney G, Gottschau A, Flierman HA, and Damsgaard MT (1990). Changing remuneration systems: effects on activity in general practice. *BMJ,* 300, 1698-1701.

Kreling DH and Mott DA (1993). The cost-effectiveness of drug utilisation review in an outpatient setting. *Pharmacoeconomics,* 4, 414-436.

Kumar S, Little P, and Britten N (2003). Why do general practitioners prescribe antibiotics for sore throat? Grounded theory interview study. *BMJ,* 326, 138-141.

Lambert BL, Salmon JW, Stubbings J, Gilomen-Study G, Valuck RJ, and Kezlarian K (1997). Factors associated with antibiotic prescribing in a managed care setting: an exploratory investigation. *Social Science & Medicine,* 45, 1767-1779.

Lancaster T, Silagy C, and Fowler G (2000). Training health professional in smoking cessation. *The Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD000214. DOI: 10.1002/14651858.CD000214.

Langley C, Faulkner A, Watkins C, Gray S, and Harvey I (1998). Use of guidelines in primary care - practitioners' perspectives. *Family Practice,* 15, 105-111.

Lawrenson R, Williams T, and Farmer R (1999). Clinical information for research: the use of general practice databases. *Journal of Public Health Medicine*, 21, 299-304.

Lassen LC and Kristense FB (1992). Peer comparison feedback to achieve rational and economical drug therapy in general practice: a controlled intervention study. *Scandinavian Journal of Primary Health Care,* 10, 76-80.

Le Grand A, Hogerzeil HV, and Haaijer-Ruskamp F (1999). Intervention research in rational use of drugs: a review. *Health Policy and Planning,* 14, 89-102.

Le Grand J (2003). *Motivation, agency and public policy: of knights and knaves, pawns and queens.* Oxford: Oxford University Press.

Levin PF (1999). Test of the Fishbein and Ajzen models as predictors of health care workers' glove use. *Research in Nursing and Health,* 22, 295-307.

Levitt B and March JG (1995). Chester I. Barnard and the intelligence of learning. In *Organizationn theory: from Chester Barnard to the present and beyond.* 2nd edition, Williamson OE (eds). pp 11-37. New York: Oxford University Press.

Lewin K (1951). *Field theory in social science.* New York: Harper Row.

Lewin SA, Skea ZC, Entwistle V, Zwarenstein M, and Dick J (2001). Interventions for providers to promote a patient-centred approach in clinical consultations. *The Cochrane Database of Systematic Reviews,* Issue 4. Art. No.: CD003267. DOI: 10.1002/14651858.CD003267.

Lewis FM (2002). Perspectives on models of interpersonal health behaviour. In *Health behavior and health education: theory, research and practice.* 3rd edition, Glanz K, Rimer BK, and Lewis FM (eds). pp 265-273. San Francisco: Jossey Bass.

Lewis G, Sharp D, Bartholomew J, and Pelosi AJ (1996). Computerized assessment of common mental disorders in primary care: effect on clinical outcome. *Family Practice,* 13, 120-126.

Leys M (2003). Health care policy: qualitative evidence and health technology assessment. *Health Policy,* 65, 217-226.

Little P, Dorward M, Warner G, Stephens K, Senior J, and Moore M (2004). Importance of patient pressure and perceived pressure and perceived medical need for investigations, referral, and prescribing in primary care: nested observational study. *BMJ,* 328, 444-446.

Lloyd DCEF, Harris CM, and Roberts DJ (1995a). Specific therapeutic group age-sex related prescribing units (STAR-PUs): weightings for analysing general practices' prescribing in England. *BMJ,* 311, 991-994.

Lloyd DCEF, Harris CM, and Clucas DW (1995b). Low income scheme index: a new deprivation scale based on prescribing in general practice. *BMJ,* 310, 165-169.

Lloyd DCEF, Roberts DJ, and Sleator D (1997). Revisions of the weights for the Age Sex Temporary Resident Originated Prescribing Unit. *British Journal of Medcial Economics,* 11, 81-85.

Lobach DF and Hammond WE (1994). Development and evaluation of a Computer-Assisted Management Protocol (CAMP): improved compliance with care guidelines for diabetes mellitus. *Proceedings - the Annual Symposium on Computer Applications in Medical Care,* 787-791.

Lobach DF and Hammond WE (1997). Computerized decision support based on a clinical practice guideline improves compliance with care standards. *American Journal of Medicine,* 102, 89-98.

Logan J and Graham ID (1998). Toward a comprehensive interdisciplinary model of health care research use. *Science Communication,* 20, 227-246.

Lohr K and Brook RH (1980). Quality of care in episodes of respiratory illness among Medicaid patients in New Mexico. *Annals of Internal Medicine,* 92, 99-106.

Lohr K and Institute of Medicine (1990). *Medicare: a strategy for quality assurance, Volume I.* Washington, DC: National Academies Press.

Lohr K and Steinwachs D (2002). Health services research: an evolving definition of the field. *Health Services Research,* 37, 7-9.

Lomas J (1991). Words without action? The production, dissemination, and impact of consensus recommendations. *Annu Rev Public Health,* 12, 41-65.

Lomas J and Haynes RB (1988). A taxonomy and critical review of tested strategies for the application of clinical practice recommendations: from "official" to "individual" clinical policy. *American Journal of Preventive Medicine,* 4, 77-94.

Long AF and Sheldon TA (1992). Enhancing effective and acceptable purchaser and provider decisions: overview and methods. *Quality in Health Care,* 1, 74-76.

Maclure M, Dormuth C, Naumann T, McCormack J, Rangno R, Whiteside C, and Wright JM (1998). Influences of educational interventions and adverse news about calcium-channel blockers on first-line prescribing of antihypertensive drugs to elderly people in British Columbia. *Lancet,* 352, 943-948.

Maibach ED, Rothschild ML, and Novelli WD (2002). Social marketing. In *Health behavior and health education: theory, research and practice.* 3rd edition, Glanz K, Rimer BK, and Lewis FM (eds). pp 437-461. San Francisco: Jossey Bass.

Mainous AG, Hueston WJ, Love MM, Evans ME, and Finger R (2000). An evaluation of statewide strategies to reduce antibiotic overuse. *Family Medicine,* 32, 22-29.

Maisonneuve H and Tiiu O (1999). From the lifecycles of clinical evidence to the learning curve of clinical experience. *Journal of Evaluation in Clinical Practice,* 5, 417-421.

Majeed A, Evans N, and Head P (1997). What can PACT data tell us about prescribing in general practice? *BMJ,* 315, 1515-1519.

Majeed A, Ferguson J, and Field J (1999). Prescribing of beta-2 agonists and inhaled steroids in England: trends between 1992 and 1998, and association with material deprivation, chronic illness and asthma mortality rates. *Journal of Public Health Medicine,* 21, 395-400.

Malekzadeh R, Mokri A, and Azarmina P (2001). Medical science and research in Iran. *Archives of Iranian Medicine,* 4, 27-39.

Malik NN (2004). Over-the-counter statins. *Lancet,* 364, 411-

Mandelblatt J and Kanetsky P (1995). Effectiveness of interventions to enhance physician screening for breast cancer. *Journal of Family Practice,* 40, 162-171.

Mann AH, Blizard R, Murray J, Smith JA, Botega N, MacDonald E, and et al (1998). An evaluation of practice nurses working with general practitioners to treat people with depression. *British Journal of General Practice,* 48, 875-879.

March JG, Sproull L, and Tamuz M (2003). Learning from samples of one or fewer. *Quality and Safety in Health Care,* 12, 465-471.

Marteau T, Sowden A, and Armstrong D (2002). Implementing research findings into practice: beyond the information deficit model. In *Getting research findings into practice.* 2nd edition, Haines A and Donald A (eds). pp 68-76. London: BMJ Books.

Mason J, Freemantle N, Nazareth I, Eccles M, Haines A, and Drummond M (2001). When is it cost-effective to change the behavior of health professionals? *JAMA,* 286, 2988-2992.

Mayer J and Piterman L (1999). The attitudes of Australian GPs to evidence-based medicine: a focus group study. *Family Practice,* 16, 627-632.

Mays N, and Pope C (1995). Rigour and qualitative research. *BMJ,* 311, 109-112.

Mays N, and Pope C (2000). Assessing quality in qualitative research. *BMJ,* 320, 50-52.

McCarty MC, Hennrikus DJ, Lando HA, and Vessey JT (2001). Nurses' attitudes concerning the delivery of brief cessation advice to hospitalized smokers. *Preventive Medicine,* 33, 674-681.

McColl A, Roderick P, Gabbay J, Smith H, and Moore M (1998). Performance indicators for primary care groups: an evidence based approach. *BMJ,* 317, 1354-1260.

McColl E, Jacoby A, Thomas A, Soutter J, Bamford C, Steen N, Thomas R, Harvey E, Garratt A, and Bond J (2001). Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technology Assessment,* 5(31),

McConnell TS, Cushing AS, Bankhurst AD, Healy JL, McIlvenna PA, and Skipper BJ (1982). Physician behavior modification using claims data: tetracycline for upper respiratory infection. *Western Journal of Medicine,* 137, 448-450.

McDonald CJ, Wilson GA, and McCabe GP (1980). Physician response to computer reminders. *JAMA,* 244, 1579-1581.

McEwen L, Farjo R, and Foxman B (2003). Antibiotic prescribing for cystitis: how well does it match published guidelines? *Annals of Epidemiology,* 13, 479-483.

McFarlane G (1984). *Alexander Fleming: the man and the myth.* London: Chatto and Windus, The Hogarth Press.

McNeil BJ, Pauker SG, Sox HC, and Tversky A (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine,* 306, 1259-1262.

McPake B, Kumaranayake L, and Normand C (2002). *Health economics: an international perspective.* London: Routledge.

McPherson K, Wennberg JE, Hovind OB, and Clifford P (1982). Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway. *New England Journal of Medicine,* 307, 1310-1313.

Meyer L (2002). Applying the theory of planned behavior: nursing students' intention to seek clinical experiences using the essential clinical behavior database. *Journal of Nursing Education,* 41, 107-116.

Meyer TJ, Van Kooten D, Marsh S, and Prochazka AV (1991). Reduction of polypharmacy by feedback to clinicians. *Journal of General Internal Medicine,* 6, 133-136.

Miettinen O, Henschke C, Pasmantier M, Smith J, Libby D, and Yankelevitz DF (2002). Mammographic screening: no reliable supporting evidence? *Lancet,* 359, 404-406.

Miles J and Shevlin M (2001). *Applying regression and correlation: a guide for students and researchers.* London: Sage.

Miles M and Huberman A (1994). *Qualitative data analysis: an expanded sourcebook.* 2nd London: Sage.

Miller W, Crabtree BF, McDaniel R, and Stange K (1998). Understanding change in primary care practice using complexity theory. *Journal of Family Practice,* 46, 369-376.

Millstein SG (1996). Utility of theories of reasoned action and planned behavior for predicting physician behavior: a prospective analysis. *Health Psychology,* 15, 398-402.

Miranda B, Canon J, Cuende N, Garrido G, Naya MT, Fernandez-Zincke E (2003). Disparities in access to liver transplantation in Spain. *Transplantation*, 76, 1398-1403.

Mitchell E and Sullivan F (2001). A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. *BMJ,* 322, 279-282.

Mittman B, Tonesk X, and Jacobson P (1992). Implementing clinical practice guidelines: social influence strategies and practitioner behavior change. *Quality Review Bulletin,* 18, 413-422.

Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, and Klassen TP (1998). Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet,* 352, 609-613.

Montano DE and Kasprzyk D (2002). The theory of reasoned action and the theory of planned behavior. In *Health behavior and health education: theory, research and practice.* 3rd edition, Glanz K, Rimer BK, and Lewis FM (eds). pp 67-98. San Francisco: Jossey Bass.

Mooney G (1995). Efficiency in health care: just health gains? *Australian Journal of Public Health,* 19, 330-335.

Morgenstern H (1995). Ecological studies in epidemiology: concepts, principles, and methods. *Annu Rev Public Health,* 16, 61-81.

Morse JM and Field PA (1996). Principles of data collection. In *Nursing research: the application of qualitative approaches.* 2 edition, pp 72-London: Chapman and Hall.

Moser C and Kalton G (1993). *Survey methods in social investigation.* 2nd ed. Dartmouth: Gower.

Moulding N, Silagy C, and Weller D (1999). A framework for effective management of change in clinical practice: dissemination and implementation of clinical practice guidelines. *Quality in Health Care,* 8, 177-183.

Mugford M, Banfield P, and O'Hanlon M (1991). Effects of feedback of information on clinical practice: a review. *BMJ,* 303, 398-402.

Muller R and Buttner P (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine,* 13, 2465-2476.

Mullins L (2002a). *Management and organisational behaviour.* 6th Harlow, England: Prentice Hall.

Mullins L (2002b). The nature of leadership. In *Management and organisational behaviour.* 6th edition, Mullins L (eds). pp 252-296. Harlow, England: Prentice Hall.

Mulrow CD (1994). Rationale for systematic review. *BMJ,* 309, 597-599.

Murchie P, Campbell NC, Ritchie LD, Simpson JA, and Thain J (2003). Secondary prevention clinics for coronary heart disease: four year follow up of a randomised controlled trial in primary care. *BMJ,* 326, 84-87.

Murphy E and Mattson B (1992). Qualitative research and family practice: a marriage made in heaven? *Family Practice*, 9, 85-91.

Murphy E, Dingwall R, Greatbatch D, Parker S, and Watson P (1998). Qualitative research methods in health technology assessment: a review of the literature. *Health Technology Assessment,* 2,

Mynors-Wallis L, Davies I, Gray A, Barbour F, and Gath D (1997). A randomised controlled trial and cost analysis of problem-solving treatment for emotional disorders given by community nurses in primary care. *British Journal of Psychiatry,* 170, 113-119.

National PCT Database (2004). www. (accessed in August 2004)

Newton-Syms FAO, Dawson PH, Cooke J, Feely M, Booth TG, Jerwood D, and Calvert RT (1992). The influence of an academic representative on prescribing by general practitioners. *British Journal of Clinincal Pharmacology,* 33, 69-73.

NHS Centre for Reviews and Dissemination (1994). Implementing clinical practice guidelines: can guidelines be used to improve clinical practice? *Effective Health Care,* Bulletin No 8:

NHS Centre for Reviews and Dissemination (1999). Getting evidence into practice. *Effective Health Care,* 5(1), 1-16.

Norman P and Conner M (1996). The role of social cognition models in predicting health behaviours: future directions. In *Predicting health behaviour: research and practice with social cognition models.* Conner M and Norman P (eds). pp 197-225. Buckingham: Open University Press.

North of England Asthma Guideline Development Group (1996). North of England evidence based guidelines development project: summary version of evidence based guideline for the primary care management of asthma in adults. *BMJ,* 312, 762-766.

North of England Evidence Based Guideline Development Project (2001). Evidence Based Clinical Practice Guideline: the primary care management of asthma in adults *BMJ,* http://www.bmj.com/asthma/ast01.htm (Accessed on 2003).

North of England Study of Standards and Performance in General Practice (1992a). Medical audit in general practice. I: Effects on doctors' clinical behaviour for common childhood conditions. *BMJ,* 304, 1480-1484.

North of England Study of Standards and Performance in General Practice (1992b). Medical audit in general practice. II: Effects on health of patients with common childhood conditions. *BMJ,* 304, 1484-1488.

Nyman JA, Akhtar MR, and Feldman R (1995). Does publishing the parameters that trigger review of Medicare claim change provider behavior? Results of the parameter release study. *Medical Care,* 33, 1022-1034.

O'Brien JA, Jacobs LM, and Pierce D (2000). Clinical practice guidelines and the cost of care: a growing alliance. *International Journal of Technology Assessment in Health Care,* 16, 1077-1091.

O'Connell DL, Henry D, and Tomlins R (1999). Randomised controlled trial of effect of feedback on general practitioners' prescribing in Australia. *BMJ,* 318, 507-511.

Ockene JK and Zapka JG (2000). Provider education to promote implementation of clinical practice guidelines. *Chest,* 118, 33S-39S.

Ogden J (2001). *Health psychology: a textbook.* 2nd Buckingham: Open University Press.

Olatunbosun OA, Edouard L, and Pierson RA (1998). Physicians' attitudes toward evidence based obstetric practice: a questionnaire survey. *BMJ,* 316, 365-366.

Olesen F, Dickinson J, and Hjortdahl P (2000). General practice - time for a new definition. *BMJ*, 320, 354-357.

Oppenheim AN (1996). *Questionnaire design, interviewing and attitude measurement.* 2nd London: Pinter Publishers.

Ovretveit J (2001). Organisational behaviour research in health care: an overview. In *Organisational behaviour and organisational studies in health care: reflections on the future.* Ashburner L (eds). pp 1-17. New York: Palgrave.

Oxman AD (1994). Checklists for review articles. *BMJ*, 309, 648-651.

Oxman AD and Flottorp S (2001). An overview of strategies to promote implementation of evidence based health care. In *Evidence based practice.* 2nd edition, Silagy C and Haines A (eds). pp 101-119. London: BMJ Publishers.

Oxman AD and Guyatt GH (1988). Guidelines for reading literature reviews. *Canadian Medical Association Journal*, 138, 697-703.

Oxman AD and Guyatt G (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology*, 44, 91-98.

Oxman AD, Thomson MA, Davis DA, and Haynes RB (1995). No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *Canadian Medical Association Journal,* 153, 1423-1431.

Packham C, Robinson J, Morris J, Richards C, Marks P, and Gray D (1999). Statin prescribing in Nottingham general practices: a cross-sectional study. *Journal of Public Health Medicine,* 21, 60-64.

Pagliari H and Kahan J (1999). Researching perceived barriers and facilitators to implementation: a coded review of studies. In *Changing professional practice: theory and practice of clinical guidelines implementation.* 1 edition, Thorsen T and Makela M (eds). pp 169-190. Dampfargevej: Danish Institue for Health Services Research and Development.

Pallant J (2001). *SPSS survival manual: a step by step guide to data analysing.* Buckingham: Open University.

Parchman ML (1995). Small area variation analysis: a tool for primary care research. *Family Medicine*, 27, 272-276.

Parker P (1994). The shape of leaders to come *Management Today,* July: 5-

Parrino TA (1989). The non-value of retrospective peer comparison feedback in continuing hospital antibiotics costs. *American Journal of Medicine,* 86, 442-448.

Paterson BL, Thorne SE, Canam C, and Jillings C (2001). *Meta-study of qualitative health research.* Thousand Oaks, California: Sage.

Pathman DE, Konard TR, Freed GL, Freeman VA, and Koch GG (1996). The awareness-to-adherence model of the steps to clinical guideline compliance. *Medical Care,* 34, 873-889.

Pendleton D and King J (2002). Values and leadership. *BMJ,* 325, 1352-1355.

Perez-Cuevas R, Guiscafré H, Munoz O, Reyes H, Tome P, Libreros V, and Gutiérrez G (1996). Improving physician prescribing patterns to treat rhinopharyngitis. Intervention strategies in two health systems of Mexico. *Social Science & Medicine,* 42, 1185-1194.

Perneger TV (1998). What's wrong with Bonferroni adjustments. *BMJ,* 316, 1236-1238.

Persaud J (2000). The Museum of Emotions, *BMJ*, 320, 1413.

Pharoah P (1996). Do counsellors in general practice change the prescribing of hypnotics and anxiolytics? *Primary Care Psychiatry,* 1, 263-264.

Phelps CE (2000). Information diffusion and best practice adoption. In *Handbook of health economics, volume 1A.* Culyer AJ and Newhouse JP (eds). pp 223-264. Amsterdam: Elsevier.

Pietro DA, Shyavitz LJ, Smith RA, and Auerbach BS (2000). Detecting and reporting medical errors: why the dilemma? *BMJ,* 320, 794-796.

Pippalla R, Riley D, and Chinburapa V (1995). Influencing the prescribing behaviour of physicians: A metaevaluation. *Journal of Clinical Pharmacy and Therapeutics,* 20, 189-198.

Plsek PE and Greenhalgh T (2001). The challenge of complexity in health care. *BMJ,* 323, 625-628.

Plsek PE and Wilson T (2001). Complexity, leadership, and management in healthcare organisations. *BMJ,* 323, 746-749.

Pocock SJ (1983). *Clinical trials: a practical approach.* Chichester: John Wiley & Sons.

Pope C, Ziebland S and Mays N (2000). Analysing qualitative data. *BMJ*, 320, 114-116.

Poplin C (2000). Productivity in primary care: work smarter, not harder. *Archives of Internal Medicine,* 160, 1231-1233.

Premaratne UN, Sterne JAC, Marks GB, Webb JR, Azima H, and Burney PGJ (1999). Clustered randomised trial of an intervention to improve the management of asthma: Greenwich asthma study. *BMJ,* 318, 1251-1255.

Prescribing Support Unit (2000). Average Daily Quantities & STAR(97)-PUs. Leeds, PSU.

Prescribing Support Unit (2002). Average Daily Quantities & STAR(01)-PUs. Leeds, PSU.

Priestley G, Watson WE, Rashidian A, Mozley C, Russell D, Wilson J, Cope J, Hart D, Kay D, Cowley K, and Pateraki J (2004). Introducing Critical Care Outreach: a ward – randomised trial of phased introduction in a general hospital. *Intensive Care Medicine,* 30, 1398-1404.

Prochaska JO and DiClemente CC (1983). Stages and processes of self-change of smoking: toward and integrative model of change. *Journal of Consulting and Clinical Psychology,* 51, 390-395.

Prochaska JO and DiClemente CC (1998). Comments, criteria, and creating better models: in response to Davidson. In *Treating addictive behaviors.* 2nd edition, Miller W and Heather N (eds). pp 39-45. New York: Plenum Press.

Prochaska JO, DiClemente CC, and Norcross J (1992). In search of how people change. applications to addictive behaviors. *American Psychologist,* 47, 1102-1114.

Prosser H, Almond S, and Walley T (2003). Influences on GPs' decision to prescribe new drugs--the importance of who says what. *Family Practice,* 20, 61-68.

Puffer S and Rashidian A (2004). Practice nurses intentions to use clinical guidelines. *Journal of Advanced Nursing,* 47, 500-509.

Quine L, Rutter D, and Arnold L (2000). Comparing the theory of planned behaviour and the health belief model: the example of safety helmet use among schoolboy cyclists. In *Understanding and changing health behaviour: from health beliefs to self-regulation.* Norman P, Abraham C, and Conner M (eds). pp 73-98. Amsterdam: Harwood Academic Publishers.

Raisch DW, Bootman JL, Larson LN, and McGhan WF (1990). Improving antiulcer agent prescribing in a health maintenance organization. *American Journal of Hospital Pharmacy,* 47, 1766-1773.

Raisch WD (1990a). A model of methods for influencing prescribing: Part I. A review of prescribing models, persuasion theories, and administrative and educational methods. *DICP, Annals of Pharmacotherapy,* 24, 417-421.

Raisch WD (1990b). A model of methods for influencing prescribing: Part II. A review of educational methods, theories of human inference, and delineation of the model. *DICP, Annals of Pharmacotherapy,* 24, 537-542.

Ramsay C, Grant AM, Wallace S, Garthwaite PH, Monk A, and Russell IT (2001). Statistical assessment of the learning curves of health technologies. *Health Technology Assessment,* 5(12),

Ramsay LE, Hag IU, Jackson PR, Yeo WW, Pickin DM, and Payne JN (1996). Targeting lipid-lowering drug therapy for primary prevention of coronary disease: an updated Sheffield table. *Lancet,* 348, 1251-1252.

Rashidian A, and Russell IT (2002a). Clinical guidelines in primary care: the complexities of changing prescribing behaviour in the NHS. A qualitative study. *International Society of Technology Assessment in Health Care (ISTAHC),* Berlin 18, 165.

Rashidian A, Russell D, and Russell IT (2002). Power calculation and sampling frame for a survey of prescribing behaviour in general practice using the theory of planned behaviour. *International Society of Technology Assessment in Health Care (ISTAHC),* Berlin 18, 165-166

Rashidian A, and Russell IT (2002b). Clinical guidelines in primary care: the views of general practitioners - A qualitative study. Fifth International Conference of Strategic Issues in Health Care Management, St Andrews, 51.

Rashidian A and Russell IT (2003). Towards better prescribing – a model for implementing clinical guidelines in Primary Care Organisations in the NHS. *Clinical Governance An International Journal,* 8, 26-32.

Rashidian A (2004a). *Appraising clinical guidelines.* The Royal College of Surgeons of England: Developing and using clinical guidelines workshop.

Rashidian A (2004b). Change management in the NHS. *Bulletin of Clinical Governance,* 4, 1-2.

Rashidian A (2004c). *Developing clinical guidelines: what make guidelines work?* University of Cambridge: Eastern Region Spring Public Health Training Workshop.

Rashidian A (2004d). *How to recommend when the evidence is lacking: the case of nutritional screening.* Royal College for Nursing and Midwifery: The 2nd NICE Systematic Reviewers' Forum.

Rashidian A, Black N and Russell IT (2005). Financial incentives and quality improvement. *Quality and Safety in Health Care,* 14, 227.

Raskin IE (1991). Medical treatment effectiveness research: a view from inside the agency for health care policy and research. *Evaluation & the Health Professionals,* 14, 187-200.

Ratanawijitrasin S, Soumerai SB and Weerasuriya K (2001). Do national medicinal drug policies and essential drug programs improve drug use? A review of experiences in developing countries. *Social Science & Medicine,* 53, 831-844.

Raven BH and Rubin JZ (1983). *Social psychology.* 2nd New York: John Wiley & Sons.

Ray WA, Blazer DG, Schaffner W, Federspiel CF, and Fink R (1986). Reducing long-term diazepam prescribing in office practice: a controlled trial of educational visits. *JAMA,* 256, 2536-2539.

Ray WA, Fink R, Schaffner W, and Federspiel CF (1985a). Improving antibiotic prescribing in outpatient practice: nonassociation of outcome with prescriber characteristics and measures of receptivity. *Medical Care,* 23, 1307-1313.

Ray WA, Schaffner W, and Federspiel CF (1985b). Persistence of improvement in antibiotic prescribing in office practice. *JAMA,* 253, 1774-1776.

Reinhardt UE (1999). The economist's model of physician behavior. *JAMA,* 281, 462-465.

Resnicow K, Braithwaite RL, Dilorio C, and Glanz K (2002). Applying theory to culturally diverse and unique populations. In *Health behavior and health education: theory, research and practice.* 3rd edition, Glanz K, Rimer BK, and Lewis FM (eds). pp 485-509. San Francisco: Jossey Bass.

Rice N, Dixon P, Lloyd DCEF, and Roberts D (2000). Derivation of a needs based capitation formula for allocating prescribing budgets to health authorities and primary care groups in England: regression analysis. *BMJ,* 320, 284-288.

Rice N and Jones A (1997). Multilevel models and health economics. *Health Economics,* 6, 561-575.

Rice N and Leyland A (1996). Multilevel models: application to health data. *Journal of Health Services Research and Policy,* 1, 154-164.

Ricordeau P, Duriex P, Weill A, Chatellier G, Vallier N, Bissery A, Fender P, and Allemand H (2003). Effect of a nationwide program of educational outreach visits to improve the processes of care for patients with type 2 diabetes. *International Journal of Technology Assessment in Health Care,* 19, 705-710.

Riemsma R, Pattenden J, Bridle C, Sowden A, Mather L, Watt I, and Walker A (2002). A systematic review of the effectiveness of interventions based on a stage-of-change approach to promote individual behaviour change. *Health Technology Assessment,* 6(24),

Rimer BK (2002). Perspectives on intrapersonal theories of health behaviour. In *Health behavior and health education: theory, research and practice.* 3rd edition, Glanz K, Rimer BK, and Lewis FM (eds). pp 144-159. San Francisco: Jossey Bass.

Ritchie J and Spencer L (1994). Qualitative data analysis for applied policy research. In *Analysing qualitative data.* Bryman A and Burgess R (eds). pp 173-194. London: Routledge.

Ritchie J, Spencer L, and O'Connor W (2003). Carrying out qualitative analysis. In *Qualitative research practice: a guide for social science students and researchers.* Ritchie J and Lewis J (eds). pp 219-262. London: Sage.

Robertson N, Baker R, and Hearnshaw H (1996). Changing the clinical behaviour of doctors: a psychological framework. *Quality in Health Care,* 5, 51-54.

Robson C (1993). *Real world research: a resource for social scientists and practitioner-researchers.* Oxford: Blackwell.

Robson MH, France R, and Bland M (1984). Clinical psychologist in primary care: controlled clinical and economic evaluation. *British Medical Journal Clinical Research,* 288, 1805-1808.

Rogers EM (1995a). *Diffusion of innovations.* 4th New York: The Free Press, Macmillan.

Rogers EM (1995b). Lessons for guidelines from the diffusion of innovations. *Joint Commission Journal on Quality Improvement,* 21, 324-328.

Roland MO, Zander LI, Evans M, Morris R, and Savage RA (1985). Evaluation of a computer assisted repeat prescribing programme in a general practice. *BMJ,* 291, 456-458.

Rosenstock I, Strecher V, and Becker M (1994). The health belief model and HIV risk behavior change. In *Preventing AIDS: theories and methods of behavioral interventions.* DiClemente RJ and Peterson J (eds). pp 5-24. New York: Plenum Press.

Rosko MD and Broyles RW (1988). *The economics of health care: a reference hankbook.* New York: Greenwood Press.

Ross-Degnan D, Soumerai SB, Goel P, Bates J, Makhulo J, Dondi N, Sutoto, Daryono A, Ferraz-Tabor L, and Hogan R (1996). A controlled trial of the impact of face-to-face persuasive education on diarrhoea treatment in pharmacies in two developing countries. *Health Policy and Planning,* 11, 308-318.

Roudot-Thoraval F, Romano P, Spaak F, Houssin D, Durand-Zaleski I (2003). Geographic disparities in access to organ transplant in France. *Transplantation,* 76, 1385-1388.

Rousseau N, McColl E, Newton J, Grimshaw J, and Eccles M (2003). Practice based, longitudinal, qualitative interview study of computerised evidence based guidelines in primary care. *BMJ,* 326, 314-318.

Royal College of General Practitioners (1985). *What sort of doctor?* London: Royal College of General Practitioners.

Rudge CJ, Fuggle SV, Burbidge KM (2003). Geographic disparities in access to organ transplantation in the United Kingdom. *Transplantation,* 76, 1395-1398.

Russell I and Gregson B (1981). Trainees' assessments of their trainers: statistical analysis. In *Occasional Paper 18: fouth national trainee conference: report,*

recommendations, and questionnaire. Ronalds C, Douglas A, Pereira Gray D and Selley P (eds). pp 73-80. London: The Royal College of General Practitioners.

Russell I, Grimshaw J, and Wilson B (1993). Scientific and methodological issues in quality assurance. *Proceedings of the Royal Society of Edinburgh,* 101, Section B: Quality assurance in medical care, 77-104.

Russell IT and Wilson BJ (1992). Audit: the third clinical science? *Quality in Health Care,* 1, 51-55.

Russell IT, Di Blasi Z, Lambert M, and Russell D (2000). Systematic reviews and meta-analyses: opportunities and threats. In *Evidence-based fertility treatment.* Templeton AA, Cooke ID, and O'Brien PMS (eds). London: RCOG Press

Ryan M (1999). A role for conjoint analysis in technology assessment in health care? *Int J Technol Assess Health Care,* 15, 443-457.

Ryan M and Farrar S (2000). Using conjoint analysis to elicit preferences for health care. *BMJ,* 320, 1530-1533.

Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, and Richardson WS (1996). Evidence based medicine: what it is and what it isn't. *BMJ,* 312, 71-72.

Saint S, Scholes D, Fihn SD, and et al (1999). The effectiveness of a clinical practice guideline for the management of presumed uncomplicated urinary tract infection in women. *American Journal of Medicine,* 106, 636-641.

Salamzadeh J, Patel MG, Wright DJ, and Chrystyn H (2001). Characteristics of general practices with different asthma prescribing *8th Iranian Researchers' Conference in Europe,* Manchester.

Salamzadeh J, Wong I, Hosker H, and Chrystyn H (2003). A cox regression analysis of covariates for asthma hospital readmission. *Journal of Asthma,* 40, 645-652.

Salisbury C, Bosanquet N, Wilkinson E, Bosanquet A, and Hasler J (1998). The implementation of evidence-based medicine in general practice prescribing. *British Journal of General Practice,* 48, 1849-1852.

Samanta A and Samanta J (2003). Legal standard of care: a shift from the traditional Bolam test. *Clinical medicine*, 3, 443-446.

Santoso B, Suryawati S, and Prawaitasari JE (1996). Small group intervention versus formal seminar for improving appropriate drug use. *Social Science & Medicine,* 42, 1163-1168.

Schaffner W, Ray WA, Federspiel CF, and Miller WO (1983). Improving antibiotic prescribing in office practice: a controlled trial of three educational methods. *JAMA,* 250, 1728-1732.

Schmidt H, Norman G, and Boshuizen HPA (1990). A cognitive perspective on medical expertise: theory and implications. *Academic Medicine,* 65, 611-621.

Schroeder SA (1987). Strategies for reducing medical costs by changing physicians' behavior: efficacy and impact on quality of care. *International Journal of Technology Assessment in Health Care,* 3, 39-50.

Schwartz G, Olsson A, Ezekowitz MD, Ganz P, and et al (2001). Effects of atorvastatin on early recurrent ischemic events in acute coronary syndromes. The MIRACL study: a randomized controlled trial. *JAMA,* 285, 1711-1718.

Schwartz RK, Soumerai SB, and Avorn J (1989). Physician motivation for nonscientific drug prescribing. *Social Science & Medicine,* 28, 577-582.

Scottish Intercollegiate Guidelines Network (1997a). *Diagnosis and management of epilepsy in adults: a national clinical guideline racommended for use in Scotland.* Edinburgh: SIGN.

Scottish Intercollegiate Guidelines Network (1997b). *Diagnosis and management of epilepsy in adults: a quick reference guide.* Edinburgh: SIGN.

Scottish Intercollegiate Guidelines Network (1998). *Primary care management of asthma: a national clinical guideline.* Edinburgh: SIGN.

Scrivener G and Lloyd DCEF (1995). Allocating census data to general practice populations: implications for study of prescribing variation at practice level. *BMJ,* 311, 163-165.

Segal R and Helper C (1982). Prescribers' beliefs and values as predictors of drug choice. *American Journal of Hospital Pharmacy,* 39, 1891-1897.

Sever P, Dahlöf B, Poulter N, Wedel H, Beevers G, Caulfield M, and et al (2003). Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial—Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial. *Lancet,* 361, 1149-1158.

Shadpour K (2003). *Health system profile of the Islamic Republic of Iran [Persian].* Tehran:

Shaneyfelt TM, Mayo-Smith MF, and Rothwangl J (1999). Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA,* 281, 1900-1905.

Shea S, DuMouchel W, and Bahamonde L (1996). A meta-analysis of 16 randomised controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting. *Journal of American Medical Informatics Association,* 3, 399-409.

Sheaff R, Rogers A, Pickard S, Marshall M, Campbell S, Sibbald B, Halliwell S, and Roland M (2003). A subtle governance: 'soft' medical leadership in English primary care. *Sociology of Health & Illness,* 25, 408-428.

Sheeran P, Norman P, and Orbell S (1999). Evidence that intentions based on attitudes better predict behaviour than intentions based on subjective norms. *European Journal of Social Psychology,* 29, 403-406.

Shekelle P, Eccles MP, Grimshaw JM, and Woolf SH (2001). When should clinical guidelines be updated? *BMJ,* 323, 155-157.

Shelley M, Croft P, Chapman S, and Pantin C (1996). Is the ratio of inhaled corticosteroid to bronchodilator a good indicator of the quality of asthma prescribing? Cross sectional study linking prescribing data to data on admissions. *BMJ,* 313, 1124-1126.

Shelley M, Croft P, Chapman S, and Pantin C (2000). Is the quality of asthma prescribing, as measured by the general practice ratio of corticosteroid to bronchodilator, associated with asthma morbidity? *Journal of Clinical Epidemiology,* 53, 1217-1221.

Shiffman J, Beer T, and Wu Y (2002). The emergence of global disease control priorities. *Health Policy and Planning,* 17, 225-234.

Shiffman RN, Liaw Y, Brandt CA, and Corb GJ (1999). Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. *Journal of American Medical Informatics Association,* 6, 104-114.

Shortell SM, Benet CL, and Byck GR (1998). Assessing the impact of continuous quality improvement on clinical practice: what will it take to accelerate progress. *Millbank Quarterly,* 76, 593-624.

Silagy C, Lancaster T, Gray S, and Fowler G (1994). Effectiveness of training health professionals to provide smoking cessation interventions: systematic review of randomised controlled trials. *Quality in Health Care,* 3, 193-198.

Skinner B (1953). *Science and human behaviour.* New York: Macmillan.

Smedslund G (2000). A pragmatic basis for judging models and theories in health psychology: the axiomatic method. *Journal of Health Psychology,* 5, 133-149.

Smith F, Singleton A, and Hilton S (1998a). General practitioners' continuing education: a review of policies, strategies and effectiveness, and their implications for the future. *British Journal of General Practice,* 48, 1689-1695.

Smith GD, Hart C, Watt G, Hole D, and Hawthorne V (1998b). Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. *Journal of Epidemiology and Community Health,* 52, 399-

Smith WR (2000). Evidence for the effectiveness of techniques to change physician behavior. *Chest,* 118, 8S-17S.

Snell JL and Buck EL (1996). Increasing cancer screening: a meta analysis. *Preventive Medicine,* 25, 702-707.

Solomon DH, Hashimoto H, Daltroy L, and et al (1998). Techniques to improve physicians uses of diagnostic tests. *JAMA,* 280, 2020-2027.

Soumerai S and Avorn J (1986). Economic and policy analysis of university-based drug 'detailing'. *Medical Care,* 24, 313-331.

Soumerai SB and Avorn J (1987). Predictors of physician prescribing change in an educational experiment to improve medication use. *Medical Care,* 25, 210-221.

Soumerai SB and Avorn J (1990). Principles of educational outreach ('academic detailing') to improve clinical decision making. *JAMA,* 263, 549-556.

Soumerai SB and Avorn J (1984). Efficacy and cost-containment in hospital pharmacotherapy: state of the art and future directions. *Milbank Memorial Fund Quarterly,* 62, 447-474.

Soumerai SB, Avorn J, Gortmaker S, and Hawley S (1987a). Effect of government and commercial warnings on reducing prescription misuse: the case of propoxyphene. *American Journal of Public Health,* 77, 1518-1523.

Soumerai SB, Avorn J, Ross-Degnan D, and Gortmaker S (1987b). Payment restrictions for prescription drugs under Medicaid: effects on therapy, cost, and equity. *New England Journal of Medicine,* 317, 550-556.

Soumerai SB, McLaughlin TJ, and Avorn J (1990). Quality assurance for drug prescribing. *Quality Assurance in Health Care,* 2, 37-58.

Soumerai SB, McLaughlin T, and Avorn J (1989). Improving drug prescribing in primary care: a critical analysis of the experimental literature. *Millbank Quarterly,* 67, 268-317.

Soumerai SB, Ross-Degnan D, and Kahn JS (1992). Effects of professional and media warnings about the association between aspirin use in children and Reye's syndrome. *Millbank Quarterly,* 70, 155-182.

Stergachis A, Fors M, Wagner EH, Sims DD, and Penna P (1987). Effect of clinical pharmacists on drug prescribing in a primary-care clinic. *American Journal of Hospital Pharmacy,* 44, 525-529.

Stevens J (1996). *Applied multivariate statistics for social sciences.* 3rd Mahway, NJ: Lawrence Erlbaum.

Stocking B (1985). *Initiatives and inertia - case studies in the NHS.* London: Nuffield Provincial Hospital Trusts.

Stocking B (1992). Promoting change in clinical care. *Quality in Health Care,* 1, 56-60.

Stone EG, Morton SC, Hulscher ME, and et al (2002). Interventions that increase use of adult immunization and cancer screening services: a meta-analysis. *Annals of Internal Medicine,* 136, 641-651.

Streiner D and Norman G (2003). *Health measurement scales: a practical guide to their development and use.* 3rd Oxford: Oxford University Press.

Stross JK and Bole GG (1980). Evaluation of a continuing education program in rheumatoid arthritis. *Arthritis & Rheumatism,* 23, 846-849.

Sullivan F and Mitchell E (1995). Has general practitioner computing made a difference to patient care? A systematic review of published reports. *BMJ,* 311, 848-852.

Sulmasy DP, Bloche G, Mitchell JM, and Hadley J (2000). Physicians' ethical beliefs about cost-control arrangements. *Archives of Internal Medicine,* 160, 649-657.

Summerskill W and Pope C (2002). 'I saw the panic rise in her eyes, and evidence-based medicine went out of the door.' An exploratory qualitative study of the barriers to secondary prevention in the management of coronary heart disease. *Family Practice,* 19, 605-610.

Sutton S (2000). A critical review of the transtheoretical model applied to smoking cessation. In *Understanding and changing health behaviour: from health beliefs to self-regulation.* Norman P, Abraham C, and Conner M (eds). pp 207-225. Amsterdam: Harwood Academic Publishers.

Tabachnick BG and Fidell LS (2001). *Using multivariate statistics.* 4th New York: HarperCollins.

Taylor K and Lane D (1998). Simulation applied to health services: opportunities for applying the system dynamics approach. *Journal of Health Services Research and Policy,* 3, 226-232.

Taylor VM, Montano DE, and Koepsell T (1994). Use of screening mammography by general internists. *Cancer Detection and Prevention,* 18, 455-462.

Terry DJ, Hogg MA, and White KM (1999). The theory of planned behaviour: self-identity, social identity and group norms. *British Journal of Social Psychology,* 38, 225-244.

The AGREE collaboration (2001). *Appraisal of Guidelines for Research and Evaluation: AGREE instrument.* London: The AGREE Collaboration.

The AGREE collaboration (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Quality and Safety in Health Care,* 12, 18-23.

Thomas L, McColl E, Cullum N, Rousseau N, and Soutter J (1999). Clinical guidelines in nursing, midwifery, and the therapies: a systematic review. *Journal of Advanced Nursing,* 30, 40-50.

Thomas L, McColl E, Cullum N, Rousseau N, Soutter J, and Steen N (1998). Effect of clinical guidelines in nursing, midwifery, and the therapies: a systematic review of evaluations. *Quality in Health Care,* 7, 183-191.

Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, and Harvey EL (1997). Educational outreach visits: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews*, Issue 4. Art. No.: CD000409. DOI: 10.1002/14651858.CD000409.

Thomson O'Brien MA, Oxman AD, Haynes RB, Davis DA, Freemantle N, and Harvey EL (1999). Local opinion leaders: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews*, Issue 1. Art. No.: CD000125. DOI: 10.1002/14651858.CD000125.

Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, and Harvey EL (2000). Audit and feedback versus alternative strategies: effects on professional practice and health care outcomes (Cochrane review - withdrawn). In *The Cochrane Library, issue 2.* Oxford: Update software.

Thomson O'Brien MA, Freemantle N, Oxman AD, Wolf F, Davis DA, and Herrin J (2001). Continuing education meetings and workshops: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews*, Issue 1. Art. No.: CD003030. DOI: 10.1002/14651858.CD003030.

Thorsen T and Makela M (1999). *Changing professional practice: theory and practice of clinical guidelines implementation.* 1 Dampfargevej: Danish Institue for Health Services Research and Development.

Timpka T, Ekstrom M, and Bjurulf P (1989). Information needs and information seeking behaviour in primary care. *Scandinavian Journal of Primary Health Care,* 7, 105-109.

Tonelli M (1998). The philosophical limits of evidence-based medicine. *Academic Medicine,* 73, 1234-1240.

Tong R (1991). The epistemology and ethics of consensus: uses and misuses of 'ethical' expertise. *Journal of Medicine and Philosophy,* 16, 409-426.

Tonkin A, Colquhoun D, Emberson J, Hague W, Keech A, Lane G, and et al (2000). Effects of pravastatin in 3260 patients with unstable angina: results from the LIPID study. *Lancet,* 356, 1871-1875.

Torgerson D, Iglesias C, and Reid D (2001). The economics of fracture prevention. In *The effective management of osteoporosis.* Barlow DH, Francis RM, and Miles A (eds). pp 111-121. London: Aesculapius Medical P.

Townsend P (1987). Deprivation. *Journal of Social Policy,* 16, 125-146.

Troyat T (1970). *Tolstoy.* London: Penguine Books.

Tukey JW (1962). The future of data analysis. *Annals of Mathematical Statistics,* 33, 1-67.

Ukoumunne OC (2002). A comparison of confidence interval methods for the intraclass correlation coefficient in cluster randomized trials. *Statistics in Medicine,* 21, 3757-3774.

Ukoumunne O, Gulliford MC, Chinn S, Sterne JAC, and Burney PGJ (1999a). Methods for evaluating area-wide and organisational-based interventions in health and health care: a systematic review. *Health Technology Assessment,* 3(5),

Ukoumunne O, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ, and Donner A (1999b). Methods in health services research: evaluation of health interventions at area and organisation level. *BMJ,* 319, 376-379.

van der Weijden T, Grol RPTM, and Knottnerus JA (1999). Feasibility of a national cholesterol guideline in daily practice: a randomized controlled trial in 20 general practices. *International Journal for Quality in Health Care,* 11, 131-137.

Vanselow NA, Donaldson MS, Yordy KD, and Institute of Medicine (1995). A new definition of primary care. *JAMA,* 273, 192-

Veninga CCM, Lagerlov P, Wahlstrom R, Muskova M, Denig P, Berkhof J, Kochen MM, and Haaijer-Ruskamp FM (1999). Evaluating an educational intervention to improve the treatment of asthma in four European countries. *American Journal of Respiratory and Critical Care Medicine,* 160, 1254-1262.

Vickers A (2003). Underpowering in randomized trials reporting a sample size calculation. *Journal of Clinical Epidemiology,* 56, 717-720.

Waddell DL (1991). The effects of continuing education on nursing practice: a meta-analysis. *Journal of Continuing Eduacation in Nursing,* 22, 113-118.

Wade OL and Hooh H (1972). Prescribing of drugs reported to cause adverse reactions. *British Journal of Preventive and Social Medicine,* 26, 205-211.

Walker AE, Grimshaw JM, and Armstrong EM (2001). Salient beliefs and intentions to prescribe antibiotics for patients with a sore throat. *British Journal of Health Psychology,* 6, 347-360.

Walker GJ, Hogerzeil HV, Sallami AO, Alwan AA, Fernando G, and Kassem FA (1990). Evaluation of rational drug prescribing in Democratic Yemen. *Social Science & Medicine,* 31, 823-828.

Walsh JME and McPhee SJ (1992). A systems model of clinical preventive care: an analysis of factors influencing patient and physician. *Health Education Quarterly,* 19, 157-175.

Walsh LJ, Wong CA, Cooper S, Guhan AR, Pringle M, and Tattersfield AE (1999). Morbidity from asthma in relation to regular treatment: a community based study. *Thorax,* 54, 296-300.

Walton R, Dovey S, Harvey E, and Freemantle N (1999). Computer support for determining drug dose: systematic review and meta-analysis. *BMJ,* 318, 984-990.

Walton RT, Harvey EL, Dovey S, and Freemantle N (2001). Computerised advice on drug dosage: effects on prescribing practice. *The Cochrane Database of Systematic Reviews,* Issue 1. Art. No.: CD002894. DOI: 10.1002/14651858.CD002894.

Wang VL, Terry P, Flynn BS, Williamson JW, Green LW, and Faden R (1979). Evaluations of continuing medical education for chronic obstructive pulmonary diseases. *Journal of Medical Education,* 54, 803-811.

Warner JO (1995). Review of prescribed treatment for children with asthma in 1990. *BMJ,* 311, 663-666.

Watkins C, Harvey I, Carthy P, Moore L, Robinson E, and Brawn R (2003). Attitudes and behaviour of general practitioners and their prescribing costs: a national cross sectional survey. *Quality and Safety in Health Care,* 12, 29-34.

Watson M, Gunnell D, and Peters T (2001). Guidelines and educational outreach visits from community pharmacists to improve prescribing in general practice: a randomised controlled trial. *Journal of Health Services Research and Policy,* 6, 207-213.

Watson PWB and Myers LB (2001). Which cognitive factors predict clinical glove use amongst nurses? *Psychology, Health and Medicine,* 6, 399-409.

Weinstein MC, Toy EL, Sandberg EA, and et al (2001). Modelling for healthcare and other policy decisions: uses, roles and validity. *Value in Health,* 4, 348-361.

Wensing M and Grol R (1994). Single and combined strategies for implementing changes in primary care: a literature review. *International Journal for Quality in Health Care,* 6, 115-132.

Wensing M, Laurant M, Hulscher M, and Grol R (1999). Methods for identifying barriers and facilitators for implementation. In *Changing professional practice: theory and practice of clinical guidelines implementation.* 1 edition, Thorsen T and Makela M (eds). pp 119-132. Dampfargevej: Danish Institue for Health Services Research and Development.

Wensing M, van der Weijden T, and Grol R (1998). Implementing guidelines and innovations in general practice: which interventions are effective? *British Journal of General Practice,* 48, 991-997.

West of Scotland Coronary Prevention Study Group (1998). Influence of pravastatin and plasma lipids on clinical events in the West of Scotland Coronary Prevention Study (WOSCOPS). *Circulation,* 97, 1440-1445.

WHO Collaborating Centre for Drug Statistics Methodology (1999). *Anatomical Therapeutic Chemical (ATC) classification index including defined daily doses (DDDs) for plain substances.* Oslo: World Health Organisation.

Wilkinson EK, McColl A, Exworthy M, Roderick P, Smith H, Moore M, and Gabbay J (2000). Reactions to the use of evidence-based performance indicators in primary care: a qualitative study. *Quality in Health Care,* 9, 166-174.

Wilkinson G, Allen P, Marshall E, Walker J, Browne W, and Mann AH (1993). The role of the practice nurse in the management of depression in general practice: treatment adherence to antidepressant medication. *Psychological Medicine,* 23, 229-237.

Williams A (1988). Health economics: the end of clinical freedom. *BMJ,* 297, 1183-1186.

Williamson OE (1995). *Organization theory: from Chester Barnard to the present and beyond.* 2nd New York: Oxford University Press.

Wilson T, Holt T, and Greenhalgh T (2001). Complexity and clinical care. *BMJ,* 323, 685-688.

Wood M, Ferlie E, and Fitzgerald L (1998). Achieving clinical behaviour change: a case of becoming indeterminate. *Social Science & Medicine,* 47, 1729-1738.

World Bank (1993). *World Development Report.* New York: Oxford University Press.

Worrall G, Chaulk P, and Freake D (1997). The effects of clinical practice guidelines on patient outcomes in primary care: a systematic review. *Canadian Medical Association Journal,* 156, 1705-1712.

Wright J, Warren E, Reeves J, Bibby J, Harrison S, Dowswell G, Russell I, and Russell D (2003). Effectivenesss of multifaceted implementation of guidelines in primary care. *Journal of Health Services Research and Policy,* 8, 142-148.

Yano EM, Fink A, Hirsch SH, Robbins AS, and Rubenstein LV (1995). Helping practices reach primary care goals: lessons from the literature. *Archives of Internal Medicine,* 155, 1146-1156.

Yeo GT, de Burgh SPH, Letton T, Shaw J, Donnelly N, Swinburn ME, and et al (1994). Educational visiting and hypnosedative prescribing in general practice. *Family Practice,* 11, 57-61.

Zazove P and Klinkman M (1998). Developing a CQI program in a family medicine department. *Journal of Quality Improvement,* 24, 391-406.

Zwarenstein M, Atkins J, Barr H, Hammick R, Koppel I and Reeves S (1999). A systematic review of interprofessional education. *Journal of Interprofessional Care,* 13, 417-424.

Zwarenstein M and Bryant W (2000). Interventions to promote collaboration between nurses and doctors. *The Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD000072. DOI: 10.1002/14651858.CD000072.

Zwarenstein M, Reeves S, Barr H, Hammick M, Koppel I and Atkins J (2000). Interprofessional education: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD002213. DOI: 10.1002/14651858.CD002213.